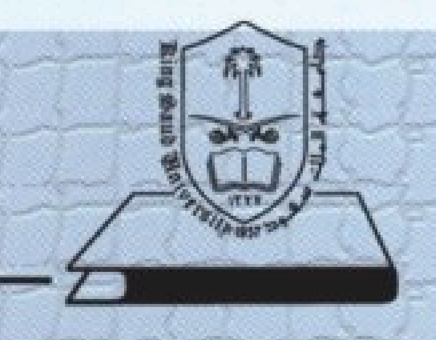
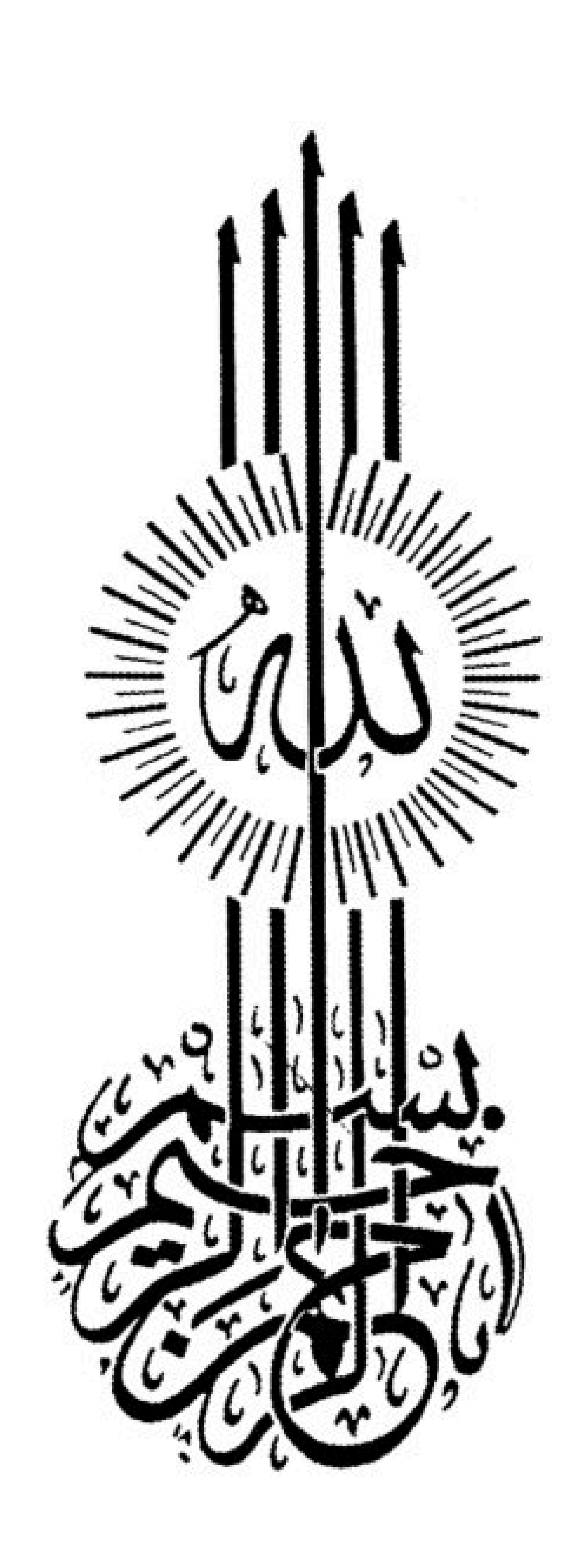
مقدمة في الإحصاء الاجتماعي

تأليف

الدكتور صالح بن محمد الصغير



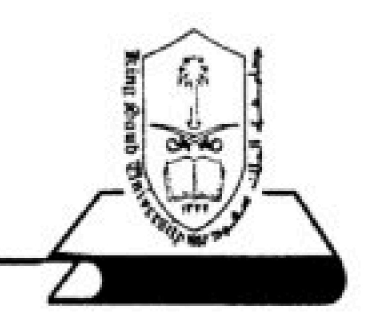
جامعة الملك سعود
النشر العلمي والمطابع


الله
بسم الله الرحمن الرحيم

مقدمة فى

الإحصاء الاجتماعي

تأليف
الدكتور صالح بن محمد الصغيّر
أستاذ مشارك - قسم الدراسات الاجتماعية
كلية الآداب - جامعة الملك سعود



النشر والمطابع - جامعة الملك سعود
ص.ب ٦٨٩٥٣ - الرياض ١١٥٣٧ - المملكة العربية السعودية

الطبعة الأولى ١٤٢٢هـ

الطبعة الثانية ١٤٢٧هـ

الطبعة الثالثة ١٤٣٢هـ

فهرسة مكتبة الملك فهد الوطنية للطباعة والنشر

الصغير، صالح بن محمد

مقدمة في الإحصاء الاجتماعي./ صالح بن محمد الصغير ط٣.-

الرياض ١٤٣٢هـ

١٨٠ص ؛ ١٧×٢٤سم

ردمك: ٥-٨١٦-٥٥-٩٩٦٠-٩٧٨

١- الإحصاء ٢- العلوم الاجتماعية- الطرق الإحصائية أ.العنوان

ديوي ٣٠١,٠١٨٢ ١٤٣٢/٤٧٣٤

رقم الإيداع: ١٤٣٢/٤٧٣٤

ردمك: ٥-٨١٦-٥٥-٩٩٦٠-٩٧٨

حكمت هذا الكتاب لجنة متخصصة شكلها المجلس العلمي بالجامعة، وقد وافق المجلس العلمي على إعادة طباعة هذا الكتاب في اجتماعه الرابع عشر للعام الدراسي ١٤٣٢/١٤٣١هـ المعقود بتاريخ ١٤٣٢/٤/٢٢هـ الموافق ٢٠١١/٣/٢٧م.



النشر العلمي والمطابع ١٤٣٢هـ

مقدمة الطبعة الثالثة

الحمد لله والصلاة والسلام على رسول الله وعلى آله وصحبه وبعد :

فقد اعتاد كثير من المؤلفين ذكر الإقبال على مؤلفاتهم ، ونفادَ الطبعات منها ، واهتمام الناس بها. وحسبي من كتابي هذا أني أول المستفيدين منه ، فإنني أعودُ له كل مرة احتاج فيها إلى معلومة إحصائية في العديد من الأبحاث التي أنجزتها.

ولقد تميز هذا الكتاب بالوضوح والبساطة في الشرح بحيث يستوعبه العالم والمتعلم ، حيث تم تبسيط مادة هذا الكتاب حتى تكون مستساغة للقارئ يتقبلها ويتفهمها ويستخدمها باستساغة وإتقان ، وفي ظني هذا هو سر النفاد السريع للطبعتين الأولى والثانية من هذا الكتاب.

ويتقدم المؤلف بخالص الشكر لجامعة الملك سعود التي أقرت هذا الكتاب ونشرته.

والله ولي التوفيق

المؤلف

مقدمة الطبعة الثانية

أصبح علم الإحصاء من العلوم التي لا غنى للباحثين في مختلف التخصصات العلمية عنه، وذلك للدور الذي يؤديه في شأن بلورة المنهج العلمي وبنائه وترسيخه في سائر الأبحاث؛ ولهذا السبب باتت تدرس المبادئ الأساسية لهذا العلم لمعظم التخصصات الطبيعية منها والإنسانية بما فيها الدراسات الاجتماعية. وهنا تكمن أهمية الأسلوب الذي تعرض من خلاله أسس ومبادئ الإحصاء الاجتماعي للمبتدئين في هذا المجال المعرفي.

ويتألف هذا الكتاب من سبعة فصول. فبعد تعريف الإحصاء كعلم، يضطلع الفصل الأول منها بتوضيح الالتحام العضوي بين الإحصاء والبحث الاجتماعي منذ الشروع في ملاحظة الظاهرة الاجتماعية وحتى اتخاذ القرار بشأنها بناء على نتيجة البحث، ولقد سمي هذا الفصل "دور الإحصاء في البحث الاجتماعي". أما الفصل الثاني "جمع البيانات" فقد عرّف ابتداء مفهوم جمع البيانات قبل أن ينطلق نحو إيراد وشرح مصادر جمع البيانات من ثانوية وأولية ثم أساليب جمع المعلومات من حيث درجة شمول الدراسة لمفردات المجتمع، أهو شمول عام (الحصر الشامل لجميع عناصر المجتمع) أم شمول جزئي (الحصر بالعينة)؟ كما وعدّد مزايا وعيوب كل من طريقتي الحصر قبل أن يذهب بعد ذلك إلى مسألة تصنيف البيانات إلى كمية (متصلة أو متقطعة) ونوعية. ثم يختتم مادته بشرح موازين قياس المتغيرات المفصلة في تحديد نوع المؤشرات الإحصائية التي ينبغي الأخذ بها عند التحليل الإحصائي وفقا لتصنيف نوع البيانات حسب تلم الموازين. ولقد خصص الفصل الثالث "عرض البيانات" لتقديم وشرح الوسائل الإيضاحية المعنية بتيسير فهم الخصائص الأساسية لمجموعة البيانات الخام المتحصل عليها بواسطة الباحث. ولقد ضمت هذه الوسائل سبل

العرض البياني للقراءات كالأعمدة والخطوط والمنحنيات والدوائر البيانية، ثم كيفية العرض الجدولي للبيانات مع شرح لأنواع الجداول الإحصائية من بسيطة ومركبة، كما ضمت كذلك شرحا وافياً لأنواع التوزيعات التكرارية لكل من نوعي البيانات الكمية والوصفية وسبل تمثيل هذه التكرارات بيانيا سواء كان ذلك بواسطة الأعمدة المتلاصقة أو المنحنيات على اختلاف أنواعها.

ولقد انصب الفصل الرابع "مقاييس النزعة المركزية" في تبيان كيفية إيجاد المتوسطات الشائعة (كالمتوسط والوسيط والمنوال) للبيانات حسابيا وبيانيا مع إيضاح مزايا وعيوب كل من تلك المتوسطات التي تضطلع بدور تلخيصي مهم لمجمل المعلومات المتحصل عليها. أما الفصل الخامس "مقاييس التشتت" فقد تناول بالشرح مقاييس التشتت الرئيسية كالانحراف المعياري والتباين والمدى ونصف المدى الربيعي مثلما تناول بالشرح أيضا الدرجة المعيارية ومعامل الاختلاف كمقياسين مهمين لمقارنة التباين وسط مجموعتين أو أكثر من القراءات المستقلة بعضها عن بعض. واهتم الفصل السادس "الارتباط" بالارتباط ومعاملاته، وتحديدا معامل ارتباط بيرسون ومعامل ارتباط سبيرمان للرتب؛ لشيوع استخدامهما في استكشاف العلاقة بين المتغيرات حيث يتناول بالشرح مفهوم الارتباط قبل أن يقدم على التطرق إلى مدى الحاجة إلى مقاييس كمية جديرة بوصف تلك العلاقة، خاصة عند تحليل البيانات الفاصلة والرتبية. وقد أحيط القارئ علما بما يحويه الفصل السابع والأخير في بداية هذا العرض، حيث يشتمل على أساليب إيجاد قيمتي مربع كاي من التكرارات الملاحظة ومن الجدول النظري لتوزيع مربع كاي، وأساليب استخدامهما لاختبار الفرضيات الصفرية في البحوث الاجتماعية. ويجدر بالذكر أنه، وعلى مدار جميع هذه الفصول، تم الركون بصفة أساسية إلى إيراد الأمثلة التطبيقية في مجال الظواهر الاجتماعية لترسيخ الشرح في أذهان الطلاب.

وتجدر الإشارة هنا أن هذه الطبعة الثانية لهذا الكتاب، ولطبيعة تعامل هذا الكتاب مع أساسيات ومبادئ علم الإحصاء في المجال الاجتماعي فأنه لا يوجد تغيرات أو إضافات تذكر عن ما ورد في الطبعة السابقة لهذا الكتاب.

المؤلف

المقدمة

أصبح علم الإحصاء القاسم المشترك الأعظم لدى الباحثين في شتى أنواع التخصصات العلمية، الطبيعية منها والإنسانية، نظرا للدور الذي يؤديه في شأن بلورة المنهج العلمي وبنائه وترسيخه في سائر الأبحاث؛ ولهذا السبب باتت تدرس المبادئ الأساسية لهذا العلم في الكليات والتخصصات المختلفة بالجامعات. ولقد شجع ذلك المهتمين في مجال الإحصاء بل ودفعهم نحو إصدار الكثير من الكتب التي تهدف إلى تقريب هذا العلم إلى أذهان الطلاب في مختلف التخصصات بما فيها تخصص علم الاجتماع رغم عدم ميل الطلاب في هذا التخصص نحو التعامل مع الرياضيات وعملياتها الحسابية. وهنا تكمن أهمية الأسلوب الذي تعرض من خلاله أسس ومبادئ الإحصاء الاجتماعي للمبتدئين في هذا المجال المعرفي. فلا بد من أن يكون أسلوب العرض ميالا نحو الشرح بالكلمات البسيطة قرينة الأمثلة التي تصب في ذات التخصص مدعمة بالرسومات الإيضاحية المصحوبة بالتغليف المبسط قبل الولوج في المعادلات الرياضية ذات الرموز التي قد تبدو معقدة أول الأمر، ليكون ذلك أدعى لاستدراج طالب الدراسات الاجتماعية نحو فهم وتفهم أفضل لأساسيات هذه المادة التي لا غنى له عنها وهو يعد في بداية الطريق نحو التطبيقات المتقدمة للعمليات الإحصائية في تحليل بيانات المسائل والمشكلات الاجتماعية.
ونحن نأمل في أن يكون هذا المؤلّف في مبادئ الإحصاء الاجتماعي قد ذهب ذلك المذهب في عرضه للمادة، خاصة وأن لدى المؤلِّف تجربة ليست بالقصيرة في تدريس

المادة نفسها لعدة سنوات متتالية في قسم الدراسات الاجتماعية بكلية الآداب بجامعة الملك سعود.

لقد صُب جلّ اهتمام هذا المؤلّف في خانة الإحصاء الوصفي كما هو متوقع من كتاب يقدم المادة للمبتدئين فيها، إلا أنه لم يهمل بعض التطبيقات الشائعة في البحوث الاجتماعية هذه الأيام لبعض المفاهيم الإحصائية المتعلقة بالإحصاء الاستدلالي، ونعني بالتحديد بعض استخدامات مربع كاي لاختبار الفروق بين سمات متغيرين اجتماعيين والذي ينفرد به الفصل السابع والأخير في هذا الكتاب. ولقد روعي في الأمثلة المنتقاة عبر جميع فصول هذا الكتاب تقريبا أن تتمحور حول الظواهر الاجتماعية، كما تم اصطحاب بعضها بوتيرة تقليدية في فصول متلاحقة حتى يتمثل الطالب كيفية تطبيقات الإحصاء الوصفي في مثال واحد مألوف لديه ولكن في مواضع متعددة ليتبين بوضوح الفوارق في هذه التطبيقات في المواضع المختلفة بصورة تكون أقرب إلى الذهن.

يتألف هذا الكتاب من سبعة فصول. فبعد تعريف الإحصاء كعلم، يضطلع الفصل الأول منها بتوضيح الالتحام العضوي بين الإحصاء والبحث الاجتماعي منذ الشروع في ملاحظة الظاهرة الاجتماعية وحتى اتخاذ القرار بشأنها بناء على نتيجة البحث، ولقد سمي هذا الفصل «دور الإحصاء في البحث الاجتماعي». أما الفصل الثاني «جمع البيانات» فقد عرّف ابتداء مفهوم جمع البيانات قبل أن ينطلق نحو إيراد وشرح مصادر جمع البيانات من ثانوية وأولية ثم أساليب جمع المعلومات من حيث درجة شمول الدراسة لمفردات المجتمع، أهو شمول عام (الحصر الشامل لجميع عناصر المجتمع) أم شمول جزئي (الحصر بالعينة)؟ كما وعدّد مزايا وعيوب كل من طريقتي الحصر قبل أن يدلف بعد ذلك إلى مسألة تصنيف البيانات إلى كمية (متصلة أو متقطعة) ونوعية. ثم يختتم مادته بشرح موازين قياس المتغيرات المفصلة في تحديد نوع المؤشرات الإحصائية التي ينبغي الأخذ بها عند التحليل الإحصائي وفقا لتصنيف نوع البيانات حسب تلك الموازين. ولقد خصص الفصل الثالث «عرض البيانات» لتقديم وشرح الوسائل الإيضاحية المعنية بتيسير فهم الخصائص الأساسية لمجموعة البيانات الخام المتحصل عليها بوساطة الباحث. ولقد ضمت

هذه الوسائل سبل العرض البياني للقراءات كالأعمدة والخطوط والمنحنيات والدوائر البيانية، ثم كيفية العرض الجدولي للبيانات مع شرح لأنواع الجداول الإحصائية من بسيطة ومركبة، كما ضمت كذلك شرحا وافيًا لأنواع التوزيعات التكرارية لكل من نوعي البيانات الكمية والوصفية وسبل تمثيل هذه التكرارات بيانيا سواء كان ذلك بوساطة الأعمدة المتلاصقة أو المنحنيات على اختلاف أنواعها.

ولقد انصب الفصل الرابع «مقاييس النزعة المركزية» في تبيان كيفية إيجاد المتوسطات الشائعة (كالمتوسط والوسيط والمنوال) للبيانات حسابيا وبيانيا مع إيضاح مزايا وعيوب كل من تلك المتوسطات التي تضطلع بدور تلخيصي مهم لمجمل المعلومات المتحصل عليها. أما الفصل الخامس «مقاييس التشتت» فقد تناول بالشرح مقاييس التشتت الرئيسة كالانحراف المعياري والتباين والمدى ونصف المدى الربيعي مثلما تناول بالشرح أيضا الدرجة المعيارية ومعامل الاختلاف كمقياسين مهمين لمقارنة التباين وسط مجموعتين أو أكثر من القراءات المستقلة بعضها عن بعض. واهتم الفصل السادس «الارتباط» بالارتباط ومعاملاته، وتحديدا معامل ارتباط بيرسون ومعامل ارتباط سبيرمان للرتب؛ لشيوع استخدامهما في استكشاف العلاقة بين المتغيرات حيث يتناول بالشرح مفهوم الارتباط قبل أن يقدم على التطرق إلى مدى الحاجة إلى مقاييس كمية جديرة بوصف تلك العلاقة، خاصة عند تحليل البيانات الفاصلة والرتبية. وقد أحيط القارئ علما بما يحويه الفصل السابع والأخير في بداية هذا العرض، حيث يشتمل على أساليب إيجاد قيمتي مربع كاي من التكرارات الملاحظة ومن الجدول النظري لتوزيع مربع كاي، وأساليب استخدامهما لاختبار الفرضيات الصفرية في البحوث الاجتماعية. ويجدر بالذكر أنه، وعلى مدار جميع هذه الفصول، تم الركون بصفة أساسية إلى إيراد الأمثلة التطبيقية في مجال الظواهر الاجتماعية لترسيخ الشرح في أذهان الطلاب. ليس هذا وحسب، وإنما جرى اختتام كل فصل من الفصول السبعة بباقة من التمارين غير المحلولة تعقبها قائمة بالاصطلاحات المهمة التي وردت بالفصل المعني وشكلت إضافة أساسية وجديدة لمعلومات الطالب المبتدئ في الإحصاء الاجتماعي. ولقد قصد من ترك التمارين بلا حل دعم النهج التربوي المتمثل في حث كل من الطالب والمدرس

الذي يضطلع بتدريس هذه المادة على النظر بجدية أكبر لعملية التفاعل بينهما وهما يحلان هذه التمارين لترسيخ مبدأ الإقناع والاقتناع.

وختاما نسأل الله التوفيق فيما يحب ويرضى، وأن ينفع أبناء الأمة الإسلامية والعربية بما قد يضيفه هذا المؤلَّف المتواضع إلى المعارف الأولية في مجال الإحصاء الاجتماعي لمن يتحلقون حول دائرة العلم من أبنائنا الطلاب وأمثالهم من المهتمين.

المؤلف

المحتويات

الفصل الثالث: عرض البيانات

الفصل الرابع: مقاييس النزعة المركزية

الفصل الخامس: مقاييس التشتت

الفصل السادس: الارتباط

الفصل السابع: دور الاختبار الإحصائي «مربع كاي» في اختبارات فروض البحوث الاجتماعية

الفصل الأول

دور الإحصاء في البحث الاجتماعي

١,١ مقدمة

يتصف علم الاجتماع – مثله مثل أي فرع من علوم المعرفة الإنسانية – بصفة التجريد. أي أنه يهتم بالخصائص المشتركة للسلوك البشري لدى جميع الحالات الفردية التي يعنى بدراستها بقصد الوصول إلى أحكام عامة تصدق على جميع المفردات. وإن أي حكم عام من هذه الأحكام يمثل عادة ما يعرف باسم القاعدة أو القانون العام. ومن أبرز شروط ومميزات هذا القانون العام هو أن يكون عالي الجاهزية لاستخدامه في تحديد ما سيحدث أو التنبؤ بما سيحدث في المستقبل. وبالطبع فإن الوصول إلى مثل هذه القاعدة أو القانون العام لا يتأتى خبط عشواء. فالباحث الاجتماعي الحصيف يشرع في ملاحظة وتتبع السلوك الإنساني الذي يهتم به في صبر وأناة ومثابرة قبل أن يقدم على مجرد افتراض شيوع مثل هذا السلوك الإنساني، أو ما يعتقد بأنه يمثل ظاهرة اجتماعية في المجتمع الذي ينوي دراسته. وفي الوقت الذي تتمكن منه الرغبة في صياغة فرضيته يجول بخاطره طيف الوسيلة التي سوف ينتهجها لاختبار صحة تلك الفرضية. ولا تعدو تلك الوسيلة أن تكون هي البوتقة التي تضم الطرق والأدوات الخاصة بالبحث المنوط به إثبات صحة أو عدم صحة تلك الفرضية. وهكذا يجد الباحث الاجتماعي نفسه في منعطف يود فيه إجراء بحث عملي إمبريقي لتمحيص صحة ودقة الحقائق المفترضة في فرضيته.

وبالطبع، فإن هذا الأمر يتطلب جمع البيانات عن الحقائق المفترضة من الميدان أولا قبل الشروع في اختبار صحة تلك الحقائق المزعومة. وهنا بالضبط يجيء دور الإحصاء الاجتماعي الذي يشكل الركيزة الأساسية في جمع المعلومات المراد الحصول

عليها وتلخيصها وشرحها واختبارها، لتكون معينة على تقويم هذه الفرضية أو تلك الفرضيات التي يعتقد أنها تساعدنا في فهم الظواهر الاجتماعية والسلوك البشري وشرحها.

١.٢ الفرضية والمتغير

لئن عرفنا فيما سبق أن الفرضيات hypotheses لا تعدو كونها أفكارا تجول بخاطر الباحثين الاجتماعيين، فقد بقي أن نعرف أن تلك الخصائص التي ندرسها اصطلح على تسميتها متغيرات variables. ولأن المتغير يمثل بهذا المفهوم الدعامة التي تبني عليها الفرضيات؛ يتعين علينا أن نميز بين مفهومي المتغير والثابت constant. ولكن قبل ذلك يجدر بنا أن نعرف على وجه الدقة ما هو المتغير في المقام الأول. يمكن تعريف المتغير بأنه «أي ظاهرة أو حدث أو خاصية تأخذ قيما تتغير من ظرف إلى آخر وتختلف هذه القيم من مفردة إلى أخرى في الكم أو الكيف». والمتغير هو الوحدة الأساسية للبيانات عند تحليلها إحصائيا. ولأن المحك العملي هو من أكثر المحكات إقناعا؛ دعنا نسوق مثالا يوضح معنى المتغير وكذلك الثابت الذي سبق وأن أرجأنا تعريفه إلى حين.

فلنفترض أن باحثا اجتماعيا زعم (أن الإناث يملن إلى إخفاء أعمارهن الحقيقية بتحيزهن نحو الأعمار الأصغر كلما تقدمن في العمر). هذه يطلق عليها فرضية، ويُستل منها متغيرا وهو (العمر) وثابتا وهو (أنثى) كمفرد للجمع (إناث) الوارد بسياق الفرضية المعنية. فعمر الأنثى يتغير (يختلف) من أنثى بعينها إلى أنثى أخرى؛ فهذه عمرها ١٥ عاما، وتلك عمرها ٢٠ عاما، . . . وهكذا. فالعمر إذن متغير بحكم تغيره من مفردة إلى مفردة أخرى. أما الأنثى فتظل هي الأنثى مهما كبر أو صغر عمرها، ولا سيما إذا وَلَدَتْ فصارت بذلك أمّا، إذ إن الأم دائما أنثى كحقيقة يعرفها كل الناس، ويتبع ذلك أنَ جنس الأم – وهو (أنثى) – ثابت دائما، أي من الثوابت التي لا تتغير في ظرفي الزمان والمكان وعبر جميع الأمهات كمفردات.

ومن أمثلة الثوابت لون الحليب فهو أبيض، وعدد أصابع اليد الواحدة فهو خمس، واتجاه البوصلة فهو شمال، ووصف الطفل فاقد والديه فهو يتيم. كما أن مستوى تعليم الأشخاص، ودخلهم المادي، وحالتهم الزواجية، وطريقة استجابتهم

للمؤثرات، كل هذه متغيرات تتغير أو تختلف باختلاف الأفراد، ويمكن للباحث الاجتماعي أن يستخدمها - مثلا - لدراسة النمو الاجتماعي واختلافاته وسط مجموعة الأطفال المولودين لأشخاص تختلف أوضاعهم الاقتصادية والاجتماعية مثل هؤلاء. إذن، فإن من أهم المميزات بين المتغير والثابت هي أن الأول يمكن استخدامه لإبراز الفوارق بين وحدات الدراسة وبالتالي التمهيد لتفسير السلوك البشري أو الظواهر الاجتماعية بناء على هذه الاختلافات أو الفوارق بين الوحدات المدروسة. أما الثاني - أي الثابت - فلا يمكن استخدامه لتحديد الفروق لأن (فاقد الشئ لا يعطيه) كما في المثل. ولنضرب بعض الأمثلة لنزيد الصورة إيضاحا.

إذا أردنا مثلا دراسة معدل الخصوبة (وهي عدد المواليد الذين ولدوا أحياء[١] لمجموعة من الأمهات) لدى مجموعة من الأمهات تم تصنيفها تبعا للعمر كمتغير، نجد أن هذا المعدل لمجموعة الأمهات في الفئة العمرية (٢٠-٢٩) سنة - منذ أن أثبتن أمومتهن - أقل من المعدل لمجموعة الأمهات نفسها عندما تصل الفئة العمرية (٣٠-٣٩) سنة. أي أن معدل الخصوبة هنا يزيد بتقدم العمر ويقل بنقصانه. أما جنس هذه المجموعة من الأمهات كثابت فلا يمكن استخدامه لدراسة الفروق في معدل الخصوبة، مثلا، لأن جنس الأم يظل أنثى سواء أزادت خصوبتها أم قلت، إذْ إن الأمهات لا يختلفن باختلاف جنسهن في أي خاصية من الخواص وماذلك ببساطة إلا لأن جنسهن ثابت، فذات الخصوبة العالية تظل أنثى وكذا ذات الخصوبة المنخفضة. كما أن لون الحليب يظل ثابتا وهو أبيض سواء كانت البقر التي درّته هي من فصيلة البقر الأمريكي أو من فصيلة البقر الهولندي؛ وبذلك لا يساعدنا لون الحليب في التمييز بين النوعين من أنواع البقر لأنه ثابت، وقس على ذلك.

ويجدر بنا أن نعيد التأكيد على أن المتغيرات والثوابت لا تنحصر فقط في خصائص الكائنات الحية، بل تشمل أيضاً الجمادات والأشياء الأخرى. ويتضح هذا من تعريف المتغير الذي أشرنا فيه إلى لفظ «مفردة» وليس «شخص» أو «فرد» حيث إن

---

(١) حسب منهج علم السكان يوصف المولود بأنه حي - وبذلك يدخل في حساب معدل الخصوبة - إذا أبدى بعد ولادته مباشرة أي حركة تدل على الحياة وإن مات بعدها مباشرة.

أول هذه الألفاظ لا يميز بين كائن حي أو جماد أو أي شيء آخر، كما أن من بين أمثلة الثوابت التي أوردناها سالفا ما لا يشير إلى خاصية كائن حي.

ولكي نمهد لتيسير فهم خاصية مهمة تحتوي عليها الفرضيات، كما سوف نتطرق إلى ذلك حالا، نقرر أنه يمكن تصنيف المتغيرات من حيث تأثير بعضها في بعضها الآخر أو تأثر بعضها ببعضها الآخر إلى مجموعتين اثنتين، إحداهما نطلق عليها نعت متغيرات مستقلة independent variables، والأخرى نسميها متغيرات تابعة dependent variables. ولا يتأتى هذا التصنيف إلا عند الزعم بوجود علاقة ما بين متغيرين على الأقل. فمثلا نقول بأن هناك علاقة بين التفكك الأسري واستشراء الجريمة في المجتمع، أو أن هناك علاقة بين مدة بقاء الوالدين مع أبنائهما ومعدل انحراف هؤلاء الأبناء، أو أن هناك علاقة بين عادة التدخين واحتمال الإصابة بسرطان الرئة. ويمكننا أن نتقدم خطوة أخرى في الزعم بتحديد صفة العلاقة في كل صيغة من الصيغ الثلاث المذكورة فنقول:

١- كلما ازداد (التفكك الأسري) في المجتمع ازداد (معدل ارتكاب الجريمة) في ذلك المجتمع.
٢- كلما زادت (مدة بقاء) الوالدين مع أبنائهما قل (معدل انحراف) هؤلاء الأبناء.
٣- كلما ازدادت (الشراهة في التدخين) ازداد (احتمال الإصابة) بمرض سرطان الرئة.

هناك ملاحظتان مهمتان تنطوي عليهما أي صيغة من الصيغ الثلاث المذكورة بعاليه - تتمثل الملاحظة الأولى في أن أي صيغة من الصيغ الثلاث هي في واقع الأمر فرضية من الفرضيات في أبسط صورها. وهي نمط من أنماط الفرضيات التي كثيرا ما يطرق أمثالها الباحثون في علم الاجتماع عند دراستهم للظواهر الاجتماعية التي يعتقدون بصحتها ويودون اختبارها. ويتضح من صيغة أي فرضية من هذه الفرضيات أن هذا النمط البسيط منها يتضمن متغيرين هما - في الصيغة الأولى - متغير التفكك الأسري ومتغير معدل ارتكاب الجريمة، مثلما أن بالصيغة الثانية متغيرين هما مدة بقاء الوالدين ومعدل الانحراف، فيما نجد بالصيغة الثالثة والأخيرة متغير الشراهة في

التدخين مع متغير احتمال الإصابة بمرض السرطان. أما الملاحظة الثانية فتتلخص في أن كل زوج من أزواج المتغيرات المتضمنة في كل فرضية يتألف من متغير نسميه متغيرا مستقلا وهو الذي يحدث أثرا ما، ومتغير تابع وهو الذي يتلقى وقع ذلك الأثر. ففي الفرضية الأولى نجد أن التفكك الأسري، وهو المتغير المستقل، يؤثر في معدل ارتكاب الجريمة، وهو المتغير التابع الذي يتلقى الأثر حيث يزيد هذا المتغير التابع بازدياد المتغير المستقل. وقس على ذلك في الفرضيتين المتبقيتين. ويتضح من هذا السياق أن المتغير المستقل هو المتغير الذي يستطيع الباحث التحكم فيه، أما المتغير التابع فهو الذي يحدث نتيجة لتحكمه هذا وهو خارج سيطرته تماما. إلا أن الباحث يتمتع بقدرته على التنبؤ بقيمة هذا المتغير التابع إذا علم قيمة المتغير المستقل التي يتعين أن تكون متاحة سلفا.

ويجب التنبيه إلى أنه بالإضافة إلى تحديد المتغيرات والفرضيات التي تتعلق بها، يتعين كذلك على الباحث الاجتماعي أن يحدد وحدة الملاحظة أو وحدة المشاهدة observation unit الخاصة ببحثه. وعادة ما يجمع الباحثون الاجتماعيون المعلومات عن أشخاص معينين، كل على حدة. فمثلا يمكن للباحث أن يجري مقابلات لتحديد ما إذا كان كبار السن يقعون ضحية للجريمة بمعدلات تفوق معدلات وقوع نظرائهم من صغار السن. ففي مثل هذه الحالة يكون المستجيب الفرد هو المعني بملاحظة الباحث الاجتماعي. إلا أن الباحثين الاجتماعيين ربما وجهوا أنظارهم نحو الوحدات الكلية aggregates بدلا من الأفراد أو الوحدات الفردية - أي أن يكون التركيز على مجاميع الناس بدلا من الأفراد الذين يؤلفون هذه المجاميع. وفي هذه الحالة يمكن للباحث، مثلا، أن يدرس العلاقة بين متوسط العمر لسكان المملكة العربية السعودية ومعدل أداء شعيرة الحج لكل منطقة من المناطق الإدارية التي تتألف منها المملكة. ففي هذه الدراسة إذًا تكون الوحدات الملاحظة أو المشاهدة (أو الوحدات تحت الدراسة) هي المناطق وليس الأفراد. وفي كلا الحالتين - أي إذا ما كانت الوحدات المدروسة هي أفراد أو مجموعات أفراد - فإن الفرضيات عادة ما تأخذ الشكل الذي يتبلور في صيغة علاقة بين متغيرين: المتغير المستقل (أي السبب المفترض) والمتغير التابع (أي الأثر المفترض) مثلما سبق التعرض إليه في أمثلة الفرضيات الثلاث.

١,٣ طرق البحث الاجتماعي

يأخذ البحث الاجتماعي أوضاعا كثيرة، ويمكن استخدامه في التحقق أو التقصي العلمي حول مختلف المسائل الاجتماعية. ومن بين أكثر طرق البحث الاجتماعي فائدة والتي قام بتوظيفها أخصائيو علم الاجتماع لاختبار فرضياتهم نذكر:

١- طريقة التجربة Experiment.
٢- طريقة المسح الاجتماعي Social survey.
٣- طريقة تحليل المضمون Content analysis.
٤- طريقة الملاحظة بالمشاركة Participant observation.

وإن من أكثر الطرق شيوعا من بين هذه الطرق الأربع هي طريقة المسح الاجتماعي والتي سوف نتطرق إليها بالتفصيل مع التركيز على أهم أداة فيها - الاستبانة - حينما نتقدم في الحديث عن جمع البيانات في الفصل القادم. أما وقد تعرضنا إلى ذكر هذه الطرق، دعنا نسوق مثالا لكل طريقة منها، وعلى نفس ترتيبها أعلاه. فمثلا، يمكن للباحث أن يجري تجربة ليقرر ما إذا كان العقاب البدني للطفل كسلوك تربوي يؤدي إلى أن يجعل ذلك الطفل أكثر عنادا في المستقبل، أو أن يجري مسحا بالعينة (الجزء المختار من مجموعة أفراد يسمى عينة) يتقصى من نتيجته اتجاهات الشباب نحو عادة التدخين، أو يجري تحليل مضمون للقيم الاجتماعية من واقع دراسة دوريات أو مجلات شبابية، أو يضطلع بالمشاركة في ملاحظة مجموعة عرقية متهمة بالتمييز العنصري.

١,٤ ما هو الإحصاء؟

لقد ذكرنا في موضوع المقدمة أن الباحث يستشعر حلول موعد استخدام الإحصاء لاختبار فرضيته بعد أن يكون قد تمكن منه الاعتقاد بصحة تلك الفرضية فعزم على اختبار صحتها. ولقد اضطررنا إلى إعطاء لمحة خاطفة عن المتغير وهو مفهوم أساسي في الإحصاء قبل التطرق إلى تعريف الإحصاء نفسه. وما ذلك إلا لمقتضيات ضرورة التعريف بالفرضية التي ينطلق منها الباحث الاجتماعي لإجراء بحثه. أما وقد اتضح ما اتضح من أمر الفرضية، فقد حق لنا الآن أن نتساءل: ما هو الإحصاء في المقام

الأول؟ ونجيب فنقرر أن كلمة «إحصاء» كاسم يشتق من الفعل المضارع «يحصي» أو الفعل الماضي «أحصى» ليست بغريبة عنا خاصة وأن اشتقاقاتها قد تشرفت بالذكر في القرآن الكريم ضمن عشرات الآيات القرآنية.[٢] ويلاحظ أن معنى الإحصاء كما يفهم من مواضع اشتقاقاته في معظم هذه الآيات لا يعدو كونه علم «العد والحصر» وهي الصورة المبسطة الأقرب إلى أذهان عامة الناس. ولقد ورد تعريف علم الإحصاء بهذه الصورة المبسطة أيضا في بعض كتب الإحصاء العتيقة الممعنة في القدَم، كما ورد في بعضها تعاريف له بأنه «علم اتخاذ القرارات» أو أنه «التعبير عن الحقيقة تعبيرا رقميا». ويشار إلى علم الإحصاء في اللغة الإنجليزية باسم statistics وهي كلمة اشتقت من كلمة state وتعني الدولة أو الولاية أو القضاء (في بعض الدول العربية)؛ ويعني الإحصاء بهذا المفهوم «علم الدولة». ولقد نسب الإحصاء إلى الدولة لأنه كان يعني جمع البيانات الخاصة بالدولة، إذ كانت البيانات التي تجمع في بادئ الأمر تتعلق بشئون الدولة وتعتبر ضرورية للحكم المستنير. ومثال ذلك الخصائص الديموغرافية للسكان مثل العدد والجنس والعمر، والبيانات المتعلقة بالصادرات والواردات والإنتاج الزراعي والحيواني. ومن مثل هذه البيانات يتمكن الحاكمون من حصر الرجال القادرين على حمل السلاح للدفاع عن الدولة، وتقدير الضرائب وغير ذلك من أمور السياسة والإدارة.

أما اليوم فقد أصبحت البيانات التي يجمعها الباحثون لا تمت في الغالب إلى شئون الدولة مباشرة كما كان الأمر في الماضي. كما اتسع استخدام الطرق الإحصائية بصورة لم تخطر على بال الذين ابتكروها حتى صارت أداة لا يستغنى عنها المتخصصون في مختلف ميادين المعرفة كالزراعة، وعلم الأحياء، والأعمال التجارية، والكيمياء، والمواصلات، والاقتصاد، والتربية، والإلكترونيات، والطب، والفيزياء، وعلم

---

(٢) من أمثلة ذلك قوله تعالى:
- ﴿ لِيَعْلَمَ أَن قَدْ أَبْلَغُوا رِسَالَاتِ رَبِّهِمْ وَأَحَاطَ بِمَا لَدَيْهِمْ وَأَحْصَىٰ كُلَّ شَيْءٍ عَدَدًا ﴾ [الجن: ٢٨].
- ﴿ ثُمَّ بَعَثْنَاهُمْ لِنَعْلَمَ أَيُّ الْحِزْبَيْنِ أَحْصَىٰ لِمَا لَبِثُوا أَمَدًا ﴾ [الكهف: ١٢].
- ﴿ وَآتَاكُم مِّن كُلِّ مَا سَأَلْتُمُوهُ وَإِن تَعُدُّوا نِعْمَتَ اللَّهِ لَا تُحْصُوهَا إِنَّ الْإِنسَانَ لَظَلُومٌ كَفَّارٌ ﴾ [إبراهيم: ٣٤].

السكان، وعلم الاجتماع، وعلم النفس، وعلم الجغرافيا، وغير ذلك من العلوم الطبيعية والإنسانية التي يصعب حصرها. ولقد تعددت لذلك الطرق والأساليب لمعالجة البيانات سواء كان ذلك من ناحية جمعها أو تلخيصها لإيضاحها أو تحليلها. ومن هذا المنطلق فإن التعريف الشامل لعلم الإحصاء يمكن أن يقرأ بأنه فرع من فروع الرياضيات، يشمل الطرق والنظريات العلمية الموجهة نحو جمع البيانات وعرضها (أي تلخيصها وتصنيفها وتبويبها بغرض وصفها) وتحليلها لاستنباط النتائج التي تستخدم لأغراض التنبؤ أو التحقق ومن ثم اتخاذ القرار حول ظاهرة معينة أو مشكلة من المشكلات.

وبالرغم من تعدد الميادين التي أصبحت تطبق فيها أسس وطرق علم الإحصاء، فإن هذه الأسس والطرق تظل في جوهرها هي نفس الأسس والطرق ولا تتغير تبعا للميدان أو المجال الذي تطبق فيه، إلا أن بعضها قد يكون أكثر ملاءمة لبعض الميادين من غيرها فيهتم الباحثون أو المتخصصون في كل مجال بالطرق التي تخدم أغراضهم. وعلى سبيل المثال، فإن المتخصصين في مجال العلوم التطبيقية (كالطب مثلا) يتطلب وضعهم معرفة ما يسمى بفن تصميم التجارب experimental designs في الإحصاء، كما أن المتخصصين في العلوم الاقتصادية والاجتماعية يتطلب وضعهم معرفة ما يطلق عليه فن تصميم الاستبانة questionnaire design لكي يحصل كل منهم على البيانات التي يطلبها بطريقة من الطرق الإحصائية العلمية. ولكن هناك مبادئ وأسس عامة في الإحصاء ينبغي أن يلم بها الجميع ولا بد لكل باحث - مهما كان تخصصه - أن يجيد معرفتها حتى يكون قادرا على فهم واستخدام الطرق الأكثر تعقيدا. هذه المبادئ والأسس العامة هي التي سوف تكون مادة هذا الكتاب فيما يلحق.

١,٥ الإحصاء الوصفي والإحصاء الاستنتاجي أو الاستدلالي

يتعين على كل مبتدئ في علوم الإحصاء أن يعرف أن الإحصاء ينقسم إلى نوعين أساسيين: إحصاء وصفي descriptive statistics وإحصاء استنتاجي inferential statistics. فالإحصاء الوصفي يختص أو يهتم بالأساليب الخاصة بتنظيم البيانات وعرضها في جداول ورسوم بيانية وأشكال هندسية، وحساب ما يطلق عليه مقاييس

النزعة المركزية measures of central tendency مثل المتوسط الحسابي والوسيط والمنوال، وكذلك مقاييس التشتت measures of dispersion مثل المدى والانحراف المعياري والتباين، وغير ذلك من المقاييس التي سوف نتطرق إليها بالتفصيل في الفصول القادمة. أما الإحصاء الاستدلالي فيهتم بالطرق التي تكشف وتستدل على المجتمع اعتمادا على ما توافر من بيانات خاصة بالعينة sample المأخوذة من هذا المجتمع، ويتناول ما يعرف بنظرية التقدير estimation theory واختبارات الفروض tests of hypotheses ومستويات الدلالة significance levels. وسوف ينصب اهتمامنا في هذا الكتاب بالدرجة الأولى على الإحصاء الوصفي.

١,٦ علاقة الإحصاء بالبحث العلمي

إذا استرجعت التعريف الوافي لعلم الإحصاء كما أوردناه في البند ٤ , ١ لوجدت أن وظائف الإحصاء تنحصر في أربع هي:

١- وظيفة جمع البيانات.

٢- وظيفة عرض وتلخيص البيانات.

٣- وظيفة تحليل البيانات واستخلاص النتائج.

٤- وظيفة اتخاذ القرار.

ونلاحظ أنه لا غنى لأي بحث من البحوث الاجتماعية عن أيّ وظيفة من هذه الوظائف. فلكي يجري الباحث بحثا يتعلق بدراسة ظاهرة معينة لا بد له من أن يجمع البيانات عنها قبل أي شئ آخر، ثم يقوم بوضع هذه البيانات في صورة يسهل معها فهمها (أي عرض البيانات)، ومن بعد ذلك يشرع في تحليلها لكي يستخلص النتائج الخاصة بهذه الظاهرة، ومن ثم يتبين أو يستنبط القوانين التي تسير تبعا لها هذه الظاهرة توطئة لاتخاذ القرار المناسب فيما يتعلق بالتدابير التي ينبغي اتباعها حيال هذه الظاهرة. ولسوف نتطرق في الفصول القادمة إلى كل وظيفة من هذه الوظائف بالتفصيل حتى يكون باكتمال الحديث عنها قد انطبع في أذهاننا تصور واضح لما ينطوي عليه الإحصاء من أهمية في الأنشطة البحثية لأخصائي علم الاجتماع.

ولئن نسبنا الإحصاء إلى علم الاجتماع فاختص عنوان هذا الكتاب بالإحصاء الاجتماعي، فما ذلك إلا لأن الطرق الإحصائية التي سوف نتناولها هي الطرق الخاصة والملائمة للتطبيق في علم الاجتماع بمسائله وظواهره التي تدور حول المجتمع البشري من سلوك وعادات وتقاليد وكسب اجتماعي. فمثلما هناك إحصاء اجتماعي، هناك أيضا إحصاء طبي وإحصاء صناعي وإحصاء زراعي، وهلم جرا. ومن هذا المنطلق، تستخدم الإحصاءات الاجتماعية كأداة لقياس درجة رفاهية الشعب، ورقي مستوى معيشته، وثقافته، وصحته، وتعليمه، وشغله، وبطالته... إلخ. فمثلا، يمكن للباحث أن يستخدم مؤشرات وإحصاءات الهجرة إلى الداخل immigration في قياس التأثير على السلوك المحلي، و إحصاءات التعليم لقياس نسبة الأمية في المجتمع، وإحصاءات المواليد والوفيات لتقدير الزيادة السكانية، وإحصاءات الزواج والطلاق لتقدير اتجاهات التماسك الأسري، كما يمكننا مقارنة إحصاءات الأجور وتقدير الثروة الوطنية مع البيانات المعروفة عن الأسعار السائدة لتقدير القوة الشرائية للسكان، وقس على ذلك ما شئت من المواضيع الاجتماعية.

١.٧ أسئلة

١- باحث اجتماعي شرع في ملاحظة موضة شبابية في المجتمع الذي يعيش فيه. متى بالضبط يستشعر حاجته إلى الإحصاء ليعينه على التقصي في الموضوع؟

٢- إذا شد انتباهك وجود بعض أفراد يمارسون التسوّل في المدينة التي تعيش فيها بيد أن مظهرهم لا يعكس موجبات هذه الممارسة، فعقدت العزم على البحث في الموضوع، أي طرق البحث الاجتماعي تتبع: المسح، أم تحليل المضمون، أم الملاحظة بالمشاركة أم التجربة؟ ولماذا؟

٣- لماذا يعتبر تعريف الإحصاء بأنه «علم اتخاذ القرار» تعريفا مبتورا؟

٤- كيف تعرّف «المتغير»؟ ولماذا يستخدم المتغير في تفسير الاختلافات بين المشاهدات بينما لا يستخدم «الثابت» لذلك الغرض؟ دعّم إجابتك بالأمثلة.

٥- عدّد أغراض الإحصاء وبين علاقة كل منها بالبحث الاجتماعي.

٦- عرّف الفرضية واشرح متى ينتهي إليها الباحث.

١,٨ اصطلاحات ينبغي تذكرها

- المتغير.
- الثابت.
- الفرضية.
- الوحدات المدروسة/ المشاهدات.
- القاعدة/ القانون العام.
- المتغير المستقل.
- المتغير التابع.

الفصل الثاني

جمع البيانات Data Collection

قبل أن يهم الباحث الاجتماعي بجمع بياناته أو معلوماته عن المشكل الاجتماعي الذي ينوي دراسته من خلال اختبار فرضيته التي قام ببنائها على النهج الذي سبق وأن تطرقنا إليه في الفصل الأول، يجب أن يكون قد وضع نصب عينيه أن قيمة نتيجة بحثه ذلك تعتمد كليا على مدى صحة المعلومات التي سوف يقوم بجمعها ودقتها. فإذا كانت البيانات التي جمعت خاطئة أو غير دقيقة فإن جميع الخطوات التي تليها كالتحليل والاستنتاج تكون كذلك خاطئة، وبالتالي تكون مضللة وكذلك الحال مع النتيجة التي يتم التوصل إليها، وبالتالي القرار الذي سوف يتخذ بشأن هذا المشكل أو تلك الظاهرة الاجتماعية. فتأمل خطورة الأمر في حالة رداءة البيانات التي يتم التعامل معها. ولكي نكون أكثر منهجية في تصور صفة البيانات التي تجمع، دعنا نصطلح على تعريف لعملية جمع البيانات بأنها «الحصول على معلومات رقمية أو وصفية تتسم بالصحة والدقة، عن ظاهرة معينة، من مصدر معين من مصادر جمع المعلومات، وفي فترة زمنية محددة. وتجمع هذه البيانات لخدمة هدف معين أو لحل مشكلة معينة». ولعل أهم عنصرين لم تتح لنا الإحاطة بشيء ما من مدلولاتهما حتى الآن من بين العناصر المذكورة في تعريف جمع البيانات المشار إليه؛ هما: مصدر جمع المعلومات، والمرجع الزمني لجمع تلك المعلومات.

٢,١ مصادر جمع المعلومات

هناك مصدران لجمع المعلومات، أحدهما يسمى المصدر الثانوي secondary source والآخر يسمى المصدر الأولي primary source. وتنسب البيانات إلى المصدر

الذي جمعت منه فنقول هذه بيانات ثانوية إذا جمعت من مصدر ثانوي، وتلك بيانات أولية إذا جمعت من مصدر أولي. فما هما إذن المصدران الثانوي والأولي؟

١.١.٢ المصدر الثانوي لجمع البيانات

قبل جمع البيانات عن أي مشكلة، يسبق ذلك دراسة وافية للمصادر التأريخية للموضوع محل البحث؛ لاحتمال توافرها كلها أو توافر جزء منها في الإحصاءات التي تنشرها مراكز البحوث أو الأجهزة الإحصائية أو الهيئات المتخصصة في الدولة، أو المنظمات العالمية مثل الأمم المتحدة، فنوفر بذلك مشقة جمعها من الميدان وما يكلفه ذلك من جهد بشري وتكاليف مادية. إذن فالمصدر الثانوي (ويسمى أحيانا المصدر التأريخي أو غير المباشر) لجمع البيانات تغلب عليه الصبغة المكتبية حيث إن الباحث لا يضطر للنزول إلى الميدان وجمع البيانات بنفسه أو بوساطة مساعديه، بل يستطيع الحصول على هذه البيانات من سجلات محفوظة تقبع منذ فترة زمنية مضت داخل هيئات، أو مؤسسات، أو وزارات، أو حتى في كتب متداولة، أو دوريات، أو من خلال شبكة المعلومات العالمية Internet إن كانت متاحة واقتضت حاجتنا الاستفادة منها. ومن أهم ما يميز البيانات الثانوية أو التأريخية عن البيانات الأولية التي سوف نتناولها بعد قليل، هو أن الأولى يقوم بجمعها في الأساس أناس آخرون لا صلة للباحث بهم إطلاقا، وإنما هو يستفيد فقط من هذه البيانات التي جمعوها منذ فترة واتفق أن صادفت حاجته الآن. والشرط الأساسي الذي يتعين على الباحث الاجتماعي التقيد به عند التعامل مع هذا النوع من المعلومات هو أن يشير إلى المصدر الذي أخذت منه بوضوح تام، ثم إلى التأريخ (على الأقل السنة سواء كانت هجرية أو ميلادية) الذي ترجع إليه هذه البيانات. فإذا كان - مثلا - لديه بيانات بعدد الحجاج الوافدين إلى مكة المكرمة في سنة معينة مصنفين تبعا لجنسياتهم، وتحصّل على هذه البيانات من وزارة الحج بالمملكة العربية السعودية؛ فإنه يتعين عليه تثبيت ذلك بوضوح أسفل الجدول الذي يحمل تلك البيانات، متضمنا اسم الوزارة كمصدر ثانوي لبياناته، والسنة التي جمعت فيها تلك البيانات. هذا بالإضافة بالطبع، إلى عنوان الجدول

الذي يجب أن يعلوه مباشرة مثلما سوف ترد الإشارة إليه عند التطرق إلى الحديث عن أنواع الجداول الإحصائية .

٢.١.٢ المصدر الأولي لجمع البيانات

ويطلق عليه أيضا اسم المصدر المباشر أو المصدر الميداني . وفيه يتم الحصول على البيانات أو المعلومات من مصادرها الأصلية عن طريق الاتصال مباشرة بوحدات أو بمفردات المجتمع الذي يُبحث ، وذلك باستخدام إحدى سبل الاتصال أو طرق البحث الاجتماعي التي تعرفنا عليها في الفصل الأول وهي : طريقة المسح الاجتماعي ، و طريقة التجربة ، وطريقة الملاحظة بالمشاركة[١] . ولأن طريقة المسح الاجتماعي هي الأكثر شيوعا واستخداما في البحوث الاجتماعية فسوف نركز عليها ونحن نتعرض بالحديث إلى المصدر الأولي لجمع المعلومات . فالمصدر الأولي يتطلب أن يقوم الباحث أو من ينوب عنه كالعداد enumerator بجمع المعلومات من المبحوث مباشرة وميدانيا . والأداة التي تستخدم لجمع البيانات من المبحوثين مباشرة تسمى الاستبانة (أو الاستمارة أو الاستبيان) وهي عبارة عن نموذج يتم تصميمه بحيث يحتوي على عدد من الأسئلة يطلب من المبحوث الإجابة عليها ، والإجابات المتحصل عليها من المبحوثين بهذه الكيفية تكون هي البيانات أو المعلومات المطلوبة . ويتم ذلك إما عن طريق المقابلة الشخصية ، أو المراسلة بالبريد ، أو المهاتفة (أو ماشابهها من وسائل الاتصال الحديثة عبر الأقمار الصناعية) . وسوف نتناول الآن كل وسيلة اتصال باستخدام الاستبانة على حدة .

٢.١.٢.١ طريقة المقابلة الشخصية Personal Interview

وفيها يقوم الباحث أو أحد مساعديه بطرح الأسئلة المعدة سلفا في الاستبانة على المبحوث ويسجل إجابات هذه الأسئلة في الاستبانة نفسها في خانات معدة خصوصا لهذا الغرض . وهناك مواصفات يجب توافرها في الاستبانة شكلا

(١) لاحظ أن طريقة تحليل المضمون لا تدخل في إطار هذا النوع من مصادر جمع المعلومات لأنها لا تحمل صفة الاتصال المباشر بوحدات المجتمع قيد البحث .

وموضوعا، وكذلك في الباحث وفي مناخ المقابلة لإنجاح هذا الأسلوب من أساليب جمع البيانات. ففيما يتعلق بالشروط المتوقع توافرها في الاستبانة الجيدة، يجب أن تكون الأسئلة الواردة بها بسيطة خالية من التعقيد، سهلة الفهم. ويجب أن يثبت في الصفحة الخارجية لغلاف الاستبانة اسم الجهة التي تتولى إجراء البحث تحت عنوان البحث مباشرة والذي يحتل وسط الصحيفة، ثم التأريخ الذي يجري فيه البحث (بالسنة والشهر واليوم ما أمكن). ومن التقاليد المتعارف عليها أن تثبت عبارة تطمئن المبحوث على أن الإفادات التي يدلي بها سوف لا تستخدم لغير الغرض العلمي تحت أي ظرف من الظروف، ويمكن أن يبلغ مضمون هذه العبارة شفاهة إلى المستجيب بالصورة التي يفهمها وتطمئنه على أن خصائصه الشخصية سوف تبقى بعيدا عن التداول الشخصي. وللإمعان في طمأنته ينبّه بأنه لا يطلب منه تسجيل اسمه ضمن البيانات التي يدلي بها.

أما فيما يخص الباحث الاجتماعي أو مساعده الذي يتولى جمع البيانات فيتعين ألا يكون غريبا عن المجتمع المعني بالدراسة، مثل أن يكلّف شخص مصري الجنسية بالطواف على الأسر بمدينة الرياض عاصمة المملكة العربية السعودية ليجمع منها بيانات عن دراسة خاصة بغلاء مهور الزواج على سبيل المثال. فينبغي في مثل هذه الحالة أن يكون جامع البيانات سعودي الجنسية على الأقل، هذا إن لم يكن سعودي الجنسية ومن المدينة نفسها. كما يجب أن يكون جامع البيانات لبقا في تعامله مع المبحوثين لكونه غريبا عليهم على أية حال، فيقدم نفسه إلى المبحوث في أدب جم مبرزا بطاقة البحث التعريفية التي يتوقع أن تحتل مكانا بارزا من صدر جلبابه، ثم يستأذن المبحوث في الجلوس في مكان ملائم - قل مثلا صالة جلوس الزائرين - ليتمكن من تسجيل إجاباته على الاستبانة على نحو مريح. وعلى جامع البيانات خاصة أن يتفادى بشكل قاطع طرح الأسئلة على المستجيب بصورة توحي إليه - أي إلى المستجيب - بإجابة معينة. مثال ذلك أن يسأل الباحث المبحوث عن عمره فيقول له: كم عمرك يا أخي. . أظن أنه في حدود ٤٠ سنة أليس كذلك؟ ففي مثل هذا الموقف، خاصة إذا كان المبحوث من رجالات البادية أو القرى الذين يغلب عليهم طابع الأمية ولا يلعب العمر دورا

رئيسا في حياتهم اليومية، ربما سارع بقبوله مقترح الباحث فأجاب بالإيجاب في حين أن الواقع قد يختلف كثيرا.

ويطبق أسلوب المقابلة الشخصية أكثر ما يطبق في المجتمعات التي تنتشر فيها الأمية كمجتمعات الدول النامية على وجه العموم. ذلك أن عدم القدرة على القراءة والكتابة يتطلب وسيطاً - بين المبحوث والاستبانة (وهو الباحث أو من ينوب عنه في هذه الحالة) - يتولى تنزيل لغة الاستبيان إلى اللغات الدارجة التي يفهمها المبحوثون الذين يغلب عليهم طابع الأمية، فيتمكنون بذلك من فهم المطلوب ويجيبون عليه كما تنبغي الإجابة الصحيحة. ومعلوم أن هذا الأسلوب من أساليب جمع البيانات تفوق تكاليفه، سواء كانت المادية أو الجهود البشرية، تكاليف أي من الأسلوبين الآخرين وهما: المراسلة بالبريد، والمهاتفة أو مثيلاتها.

٢,٢,١,٢ طريقة المراسلة بالبريد Mailing Method

تتطلب هذه الطريقة أيضا استخدام الاستبانة ذات المواصفات المحددة مع فارق أساسي عند التعامل مع المبحوث حيث سترسل إليه بالبريد على عنوانه البريدي ويتولى هو كليا الإجابة على جميع الأسئلة الواردة بها، ثم يقوم بإعادتها إلى الجهة التي تضطلع بالبحث. ويتضمن مثل هذا النوع من الاستمارات إرشادات إضافية لكيفية الإجابة وكيفية ملء الخانات المخصصة للإجابات؛ لأن المستجيب لا يتعامل سوى مع هذه الإرشادات كبديل للباحث شخصيا في الاستفسار عن أي غموض قد يكتنف الأسئلة الواردة بالاستبيان، أو النهج الذي ينبغي أن تكون عليه الإجابة. وزيادة في الحرص على استخلاص الأجوبة الصحيحة، يجب أن تكون الأسئلة متناهية الدقة والتحديد والوضوح بحيث لا تحتمل أي لبس، مع أن احتمال هذا اللبس مهما كان ضئيلا يجب أن ينتفي تماما باطلاع المبحوث على الإرشادات المصاحبة للاستبانة. وقبل الإقدام على إدخال هذه الاستبانة في مظروف وإرسالها بالبريد إلى المبحوثين، يتعين التأكد من أن هذا المظروف يحتوي - بالإضافة إلى الاستبانة - على ما يلي:

١- طابع بريدي لاستخدامه في إرجاع الاستبانة المستوفاة.

٢- السند القانوني الذي يتيح إجراء البحث ويستوجب الاستجابة.
٣- قائمة بالإرشادات التفصيلية لكيفية ملء الاستمارة.
٤- عنوان واضح للجهة التي تستقبل الاستمارات المستوفاة.
٥- المدة القصوى التي ينبغي ألا يتعداها موعد استلام الاستمارات المستوفاة.
٦- عنوان واضح للمبحوث.

وربما لاحظ القارئ أن طريقة المراسلة بالبريد تتطلب وعيا عاليا من قبل أفراد المجتمع الذي تطبق فيه الاستمارة على هذا النحو، ليس فقط على مستوى كيفية ملء الاستبانة ولكن أيضا على مستوى استشعار أهمية البحث وفائدته العلمية التي يجني ثمارها بالطبع المجتمع المعني؛ لكونه مرشحا لمستوى رفاهية أعلى باستفادته من تطبيق التوصيات المبنية على نتائج البحث. والهدف المباشر بالطبع من وراء تطبيق تلك التوصيات يصب في عملية توفير أسباب رفاهية ذلك المجتمع وتقدمه. ولن يدرك هذا المستوى من الفهم إلا شخص متعلم بشكل جيد على أقل تقدير. ولهذا، فإن هذا الأسلوب من أساليب جمع البيانات يكون مناسبا في مجتمعات الدول المتقدمة مثل دول أوروبا والولايات المتحدة الأمريكية. كما أن تطبيقه يتطلب أن يكون المجتمع الذي يطبق فيه متمتعا ببنيات أساسية متينة وشبكة اتصالات متطورة بما فيها وسائل الاتصال البريدي المختلفة (بريد عادي، أو إلكتروني بما فيه الفاكس والإنترنت) وسبل مواصلات حديثة كالطائرات والقطارات السريعة... إلخ. وباختصار، فإن وسيلة المراسلة بالبريد تتطلب درجة عالية من العمران بما فيها الطرق والكباري والمنشآت المدنية كافة بما فيها المباني ذات التشييد الجيد التي تتمتع بالإمداد الكهربائي والمائي إضافة إلى ما سبق ذكره من سبل الاتصال والمواصلات الجيدة. وهذا لا يتوافر بدرجة معقولة في الدول النامية مما يعد أحد الأسباب الرئيسة التي تقعد بها عند الأخذ بهذه الوسيلة لجمع المعلومات رغم أنها أقل تكلفة مقارنة مع وسيلة المقابلة الشخصية وهذه من أهم ميزاتها، إلا أن عيوبها تتمثل في إمكان ضياع جزء من الاستبانات خلال عمليتي الإرسال والاستلام، وإمكان عدم الاهتمام باستيفاء الاستبانات وإرجاعها في الوقت المحدد من قبل بعض المبحوثين.

٢,١,٢,٣ طريقة المهاتفة

تتميز هذه الطريقة بالسرعة في جمع البيانات لكنها ولأسباب عملية لا تطبق إلا في نوع معين من أنواع البحوث المحدودة مثل بحوث استطلاعات الرأي. ذلك أن البحوث الكبيرة - أي التي تجري على نطاق واسع وتشمل بيانات تفصيلية عن وحدات البحث - لا يمكن إجراؤها باستخدام الهاتف الذي لا يتيح للمفحوص حريته في الإجابة على البنود ذات الطابع الشخصي، إذ يكون في شك من ضمان سرية بياناته الشخصية عبر هذه الأداة. زد على ذلك أن عملية إلقاء أسئلة الاستبانة على المستجيب بهذه الكيفية وانتظار الإجابة عليها عبر نفس الوسيلة تكون غير عملية لما تستغرقه من وقت طويل وما يجره ذلك من تكاليف باهظة. إذن أنسب أنواع البحوث التي تتناسب معها هذه الوسيلة هي تلك التي لا تستغرق وقتا طويلا وتكون الأسئلة فيها مختصرة ومحددة. ومن أمثلة ذلك استطلاع رأي شريحة من المستهلكين عن أصناف تنافسية لبضائع معينة، أو استطلاع رأي جمهور الشباب في مدينة معينة عن مدى تأييدهم أو رفضهم لآراء شخصية عامة استضافتها دار التلفاز، مثلا، عن مصالح تهم القطاع الشبابي من القطاعات العمرية للسكان، وقس على ذلك.

٢,٢ أساليب جمع البيانات من حيث درجة شمولها لمفردات مجتمع البحث

بحسب مقدار تغطية مفردات مجتمع البحث، هناك أسلوبان أساسيان من أساليب جمع البيانات ميدانيا:

(أ) أسلوب الحصر الشامل Complete coverage.

(ب) أسلوب الحصر بالعينة Sample coverage.

يتم في أسلوب الحصر الشامل أخذ المعلومات أو البيانات من جميع العناصر المكونة لمجتمع البحث مثلما هو الحال - مثلا - عند إجراء تعداد سكاني قومي عام يشمل (جميع الأسر) بدولة ما، أو بحث اجتماعي يشمل (جميع الأمهات) بمدينة ما، أو بحث اجتماعي آخر يشمل (جميع الأندية الرياضية) بمنطقة ما من المناطق الإدارية المكونة لدولة من الدول. أما في أسلوب الحصر بالعينة فيتم أخذ المعلومات فقط من (بعض) مفردات مجتمع البحث على أن يكون ذلك البعض ممثلا تمثيلا صادقا

للمجتمع الذي يُدرّس ؛ ومثال ذلك ، أن يتم إجراء بحث سكاني يشمل (جزءاً من الأسر) بدولة ما أو بحث اجتماعي يشمل (جزءاً من الأمهات) بمدينة ما أو بحث اجتماعي آخر يشمل (جزءاً من الأندية الرياضية) بمنطقة ما من المناطق الإدارية المكونة لدولة من الدول .

وقبل الخوض في خصائص كل أسلوب من هذين الأسلوبين على حدة ، يتعين علينا الإشارة إلى أن عملية اختيار الأسلوب الأنسب لجمع البيانات من بين هذين الأسلوبين يتم تحديدها أول وهلة للشروع في تصميم البحث الاجتماعي . ذلك أن اختيار أسلوب جمع البيانات الملائم للبحث يتوقف على عدة محكات منها الإمكانات المادية والبشرية المتاحة لإجراء البحث ، والوقت المحدد والسرعة المطلوبة للحصول على النتائج ، وطبيعة الظاهرة المراد دراستها ودرجة تجانسها . كما أن النتائج التي نحصل عليها من تحليل بيانات الدراسة قد تختلف باختلاف الأسلوب المتبع في جمع هذه البيانات .

١,٢,٢ أسلوب الحصر الشامل

وهو ، كما أسلفنا ، الأسلوب الذي يدرس فيه الباحث جميع عناصر أو وحدات المجتمع المراد بحثه دون استثناء . والشائع أن يتبع هذا الأسلوب عندما يراد الحصول على معلومات تفصيلية عن جميع مفردات المجتمع كما هو الحال في الأبحاث القومية الكبيرة التي تجرى لأغراض التخطيط القومي الشامل ، سواء كان متوسط المدى أو بعيد المدى ، مثال ذلك : التعداد السكاني القومي[٢] ، وتعداد المنشآت الزراعية ، والصناعية ، والاجتماعية . . . إلخ . ولا يخفى أن مثل هذا النوع من الأبحاث يحتاج إلى فترة زمنية طويلة نسبيا ، وإلى كادر إداري وآخر فني مدربين تدريبا عاليا على عمليات الإدارة والحصر في الميدان ، وإلى وسائل اتصال ومواصلات ذات كفاءة عالية علاوة على

---

(٢) توصي منظمة الأمم المتحدة أن تجرى التعدادات السكانية كل خمس أو عشر سنوات كحد أقصى في كل دولة ، وذلك لأجل الحصول على بيانات اقتصادية - اجتماعية مقارنة يعول عليها في الأبعاد التخطيطية .

كبر حجمها، وفي المحصلة يحتاج هذا الأسلوب إلى مبالغ طائلة من الأموال إذا ما تقرر استخدامه لجمع البيانات. ولا يقتصر هذا الأسلوب بالطبع على مثل هذا النوع من الأبحاث الذي يعني بجمع البيانات من جميع وحدات الدراسة على مستوى القطر كله. بل يمكن أن توصف به عملية جمع المعلومات من جميع الطلاب المسجلين بجامعة الملك سعود، مثلا، بقصد دراسة اتجاهات هؤلاء الطلاب نحو ظاهرة زواج الأقارب على سبيل المثال. أو جمع المعلومات من جميع شباب مدينة الرياض الذين تنحصر أعمارهم بين ١٥ إلى ٢٥ سنة، مثلا، بهدف دراسة اتجاهات هؤلاء الشباب نحو ظاهرة «التفحيط» أو ما يعاصرها من مستجدات الظواهر بهذه المدينة. ففي كلا المثالين تشمل الدراسة جميع وحدات المجتمع المراد دراسته وليس جزءاً منها. أي أننا لم نختر عددا محدودا من طلاب جامعة الملك سعود لإجراء الدراسة عليهم كعينة ممثلة لجميع طلابها، كما أننا لم نختر عددا محدودا من شباب مدينة الرياض الذين تتراوح أعمارهم بين ١٥ و٢٥ سنة لنأخذ عنهم المعلومات عن الظاهرة المدروسة لكونهم عينة تمثل الشريحة المعنية من الشباب كمجتمع للدراسة.

٢,٢,٢ أسلوب الحصر بالعينة

العينة ببساطة هي جزء من مجتمع الدراسة كما أسلفنا. ولكن لكي يكون هذا الجزء ممثلا تمثيلا صادقا لمجتمع الدراسة، وبالتالي يكون ممكنا تعميم النتائج المتحصل عليها من دراسة هذا الجزء فقط على جميع مفردات مجتمع الدراسة، يجب أن تتصف هذه العينة بصفة ما نطلق عليه العينة الاحتمالية Probabilistic sample. وعلى نقيض العينة الاحتمالية هناك عينة غير احتمالية Non probabilistic sample. فالعينة الاحتمالية سميت هكذا لأنها تعتمد على نظرية الاحتمالات Probability theory. وأهم ما يميز العينة الاحتمالية هو أنها عينة معروف احتمالات اختيار مفرداتها مسبقاً، كما أن لكل فرد من أفراد المجتمع الذي سحبت منه العينة احتمال للاختيار ضمن أفرادها، وهذا الاحتمال يمكن حسابه ولا يساوي صفرا. ولهذا، وإذا كان أسلوب المعاينة sampling method الذي اتبع لاختيار العينة الاحتمالية أسلوبا منهجيا علميا صحيحا، فإن هاتيك العينة الاحتمالية تكون ممثلة لجميع خصائص المجتمع الذي

سحبت منه، وبالتالي يمكن تعميم نتائجها على ذلك المجتمع بعد السماح بهامش خطأ معين وفقا لنظرية الافتراضات وما يصاحبها من اختيار مستويات للدلالة مما سوف يرد ذكره بالتفصيل في فصول لاحقة. ومن أهم أنواع العينات الاحتمالية ما يلي:

(أ) العينة العشوائية البسيطة Simple random sample.

(ب) العينة العشوائية المنتظمة Systematic random sample.

(جـ) العينة العشوائية الطبقية Stratified random sample.

(د) العينة العشوائية العنقودية Clustered random sample.

ولكل نوع من هذه الأنواع طريقة معينة يتم بها سحب مفردات العينة من المجتمع المراد دراسته. ولأن طبيعة هذا الكتاب لا تستوعب متسعا لشرح طروحات الاختيار تلك، يمكن للمهتمين الرجوع إلى الكتب المنهجية في علم العينات[٣] لسبر غورها تفصيلا.

أما العينات غير الاحتمالية فهي التي لا تنبني أصلا على نظرية الاحتمالات، إذ إنه ليس بإمكاننا معرفة أية احتمالات للاختيارات مسبقا. وبذلك فهي عينات غير احتمالية - أي أنه ليس بمستطاعنا تعميم نتائجها على المجتمع الذي تم سحبها منه. ومن أهم أنواع العينات غير الاحتمالية:

(أ) العينة بالمصادفة (أو العينة الصدفية) Accidental sample.

(ب) العينة الحصصية (أو التدرجية) Quota sample.

(جـ) العينة العمدية (أو المقصودة) Purposive sample.

ولبساطة اختيار المفردات وفقا لهذه العينات غير الاحتمالية نعرض لكل واحدة منها في إيجاز[٤]. ففي العينة بالمصادفة لا يتم اختيار المفردات طبقا لأي نسق وإنما يتم

---

(٣) انظر مثلا:

1. Kish, "Survey Sampling", 1965.

2. Hansen, Hurwitz, and Madow, "Sample Survey Methods and Theory", 1953.

(٤) انظر مثلا: الشربيني، الإحصاء اللابرامتري، ١٩٩٠م، ص١٩، ٢٠.

اعتماد من يتحصل عليه صدفة، أو من يشارك من الأفراد متطوعا في البحث. مثال ذلك أن يتم اختيار أول عشرين مرتادا لمكتبة الملك فهد بالرياض في صبيحة يوم بعينه. أما في العينة الحصصية فيتم اختيار المفردات وفقا لنسب مئوية تتناسب ونسبة وجود خاصية مجتمع تلك المفردات في المجتمع الذي سحبت منه. فمثلا إذا أردنا أن نختار عينة حصصية من مجتمع تم تصنيفه تبعا لخاصية المهنة: عامل، ومزارع، ومهندس... إلخ، نقوم أولا بتحديد النسبة المئوية لشاغلي كل مهنة من مجموع شاغلي جميع المهن. بعد ذلك نحدد نسبة مئوية يتم سحبها من كل مهنة وفقا لوزن هذه المهنة بالنسبة إلى الحجم الكلي لشاغلي المهن بالمجتمع المعني حتى نضمن تمثيل كل مهنة في العينة تبعا لحجمها في مجتمع الدراسة. وفيما يتعلق بالعينة العمدية أو المقصودة فإن اختيار المفردات يتم بناء على خبرة الباحث ومن ثم قناعته بأن المفردات التي تم اختيارها في العينة العمدية تمثل المجتمع الذي سحبت منه دونما تحيز لشريحة فيه على حساب الأخرى. ولنفي احتمال التحيز هذا، يتعين على الباحث تدوين مبرراته بعدم وجود احتمال مثل هذا التحيز لتكون في متناول من بنيته تقويم نتائج البحث المتحصل عليها من عينة تم اختيارها على هذا المنوال. ويعتبر مثالا للعينة المقصودة أن يعمد باحث إلى اختيار بعض من الأحياء السكنية بمدينة «جدة» بالمملكة العربية السعودية زاعما أنها تمثل مختلف الطبقات الاجتماعية بالمدينة ليجري بذلك بحثا عن اتجاه سكان هذه المدينة بمختلف طبقاتها الاجتماعية نحو ظاهرة اجتماعية مستجدة.

٣.٢.٢ ملخص مزايا وعيوب كل من أسلوبي الحصر الشامل والحصر بالعينة

يمتاز أسلوب الحصر الشامل بتوافره على إمكانية الحصول على معلومات تفصيلية عن جميع مفردات المجتمع مما يتيح أرضية مقنعة للتخطيط الاجتماعي و/ أو الاقتصادي بعيدا عن احتمالات التكهن. إلا أن عيوبه تتمثل في أنه يتطلب أموالا طائلة وتكاليف مرتفعة مادية وبشرية، كما أنه يتطلب وقتا طويلا لجمع المعلومات. ونقيض هذين الأمرين يمثل مزيتين مهمتين لأسلوب الحصر بالعينة الذي يتطلب أموالا أقل ووقتا أقصر نسبيا لجمع المعلومات نظرا لمحدودية درجة شمول حصر مفردات المجتمع المعني بالدراسة. وإن بدا للقارئ بأن حقيقة أخذ المعلومات عن عدد محدود

من أفراد المجتمع عند الأخذ بأسلوب الحصر بالعينة يمثل عيبا لهذا الأسلوب فإن هذا العيب ينتفي تماما إذا ما اتبع النهج العلمي الدقيق في اختيار العينة على أن تكون هذه العينة عينة احتمالية . كما أن أسلوب الحصر بالعينة يمتاز بالسرعة في إظهار النتائج ، وهذه الأخيرة تمتاز بالدقة نظرا لدقة البيانات المجموعة بهذا الأسلوب مقارنة مع نظيرتها التي جمعت عن طريق أسلوب الحصر الشامل . زد على ذلك أنه ، وفي بعض الأحوال ، يجب التعويل على أسلوب الحصر بالعينة بدلا من نظيره الحصر الشامل مثل الحالات التي يصعب فيها فحص المجتمع بكامله أو حتى يستحيل في حالات أخرى . أمثلة ذلك : حصر وفحص الثروة السمكية بالخليج العربي ، أو فحص الإنتاج اليومي لمصنع ما ، أو حصر المنحرفين بدولة ما أو حتى بمدينة ما ، أو فحص الأدوية والأطعمة ، أو فحص دم المريض . . وقس على ذلك .

ويمكن القول إجمالا أن أسلوب الحصر بالعينة يمتاز على أسلوب الحصر الشامل بالمزايا التالية :

(أ) قلة التكاليف .

(ب) السرعة في إظهار النتائج .

(جـ) دقة البيانات .

(د) الخيار الوحيد في حالة صعوبة أو استحالة فحص المجتمع بكامله .

٤ . ٢ . ٢ مفاهيم تتعلق بالبيانات التي يتم جمعها

يجب على الطالب أن يدرك أولا أن «المجتمع الإحصائي» الذي يجمع الباحث - أي باحث - بياناته من جميع مفرداته أو من بعض منها لا يتطابق مفاهيميا مع «المجتمع» الذي يرمي إليه الباحث الاجتماعي والذي يعني المجتمع البشري . ففي علم الإحصاء يعني المجتمع population أي مجموعة متجانسة نوعيا سواء كانت مجموعة بشر mankind أو مجموعة حيوانات animals أو مجموعة أحياء أخرى other living creatures أو مجموعة أشياء objects كالحجارة أو الأتربة أو الغازات . وكل مجموعة من هذه المجموعات تسمى مجتمعا population في مفهوم علم الإحصاء وخصوصا علم العينات sampling . وإذا حصرنا التركيز على المجتمعات

الإنسانية باعتبارها أيضا مجتمعات إحصائية ، فإن البيانات التي نقوم بجمعها من مفردات هذه المجتمعات (أي ، من الناس) ما هي إلا خصائص هذه المفردات سواء كانت هذه الخصائص خصائص ديموغرافية كالعمر والجنس ، أو فيزيائية كالطول والوزن ، أو اجتماعية كالمستوى التعليمي والحالة الزواجية ، أو اقتصادية كالدخل الشهري وحجم الأملاك . ومادام أن هذه الخصائص قدرها أن تختلف أو تتباين أو تتغير من مفردة (أي ، شخص) إلى مفردة أخرى فقد اصطلح على تسميتها متغيرات variables . وهكذا ننتهي إلى أن البيانات التي نقوم بجمعها من الميدان ما هي إلا مجموعة متغيرات . ولقد أشرنا في الفصل الأول إلى أن المتغير هو الوحدة الأساسية أو اللبنة القاعدية التي ينبني عليها تحليل البيانات . كما تطرقنا في الفصل نفسه إلى تصنيف المتغيرات إلى متغيرات مستقلة وأخرى تابعة ، إلا أن للمتغيرات أوجه تصنيف أخرى نتعرض لها الآن قبل أن نخوض في عملية القياس measurement التي تخضع لها هذه المتغيرات قبل أن تعطينا إشارة البدء في عملية تحليلها توطئة لاستخلاص النتائج .

١ ,٤ ,٢ ,٢ التصنيف طبقا لما تعبر عنه قيمة المتغير من مقدار أو صفة

يوصف المتغير بأنه متغير كمي إذا كانت قيمته تشير إلى مقدار quantity ما لدى المفردة من خاصية (مثل الطول ، والوزن ، والعمر ، والدخل . . . وهلم جرا) . وطالما أن قيمة المتغير هنا تحمل معنى كميا ، يمكن بذلك ترتيب المفردات طبقا لهذا النوع من الخصائص (أي التي يمكن التعبير عنها كميا) من الأكبر إلى الأصغر والعكس بالعكس .

وإذا كانت قيمة المتغير لا تعبر عن مقدار الخاصية عند المفردة بل تعبر عن وجود هذه الخاصية أو عدم وجودها لدى المفردة - أي إذا كانت المفردة تمتلك هذه الخاصية أو لا تمتلكها - مثل الجنس (ذكر ، أو أنثى) ، التخصص (علمي ، أو أدبي) ، الجنسية (مصري ، أو أردني ، أو سعودي . . . إلخ) ، اللون (أحمر ، أو أبيض ، أو أزرق . . . إلخ) فإن الخاصية تحمل معنى نوعيا qualitative ويسمى المتغير عندئذ متغيرا نوعيا . ويلاحظ أنه لا يمكن ترتيب المفردات طبقا لهذه الخاصية من الأكبر إلى الأصغر أو العكس .

٢,٢,٤,٢ تصنيف المتغيرات الكمية طبقا للطبيعة الرقمية لقيمتها: صحيحة أم كسرية

تصنف المتغيرات الكمية من حيث الطبيعة الرقمية لقيمها إلى :

(أ) متغيرات كمية متصلة Continuous variables .

(ب) متغيرات كمية منفصلة Discrete variables .

فالمتغيرات الكمية المتصلة تأخذ أي قيمة سالبة أو موجبة، صحيحة أو كسرية، مثل درجات الحرارة، والأعمار، والأطوال . أما المتغيرات الكمية المنفصلة فتأخذ قيما صحيحة فقط لا تحتمل الكسور مثل عدد أفراد الأسرة، والفصل الدراسي، وعدد الكراسي .

٢,٢,٤,٣ التصنيف حسب ميزان القياس

تصنف المتغيرات تبعا لمستويات أو لموازين القياس التي تقاس عليها للتمكن من اختيار نوع التحليل الإحصائي المناسب لها إلى أربعة أنواع رئيسة هي :

(أ) متغيرات اسمية Nominal variables .

(ب) متغيرات رتبية Ordinal variables .

(جـ) متغيرات فاصلة (أو فترية) Interval variables .

(د) متغيرات نسبية Ratio variables .

وسوف يجري فهم هذه الأنواع تلقائيا بعد الاطلاع على شرح قياس المتغيرات وموازين هذا القياس فيما يلي .

٢,٢,٥ موازين قياس المتغيرات Variable Measurement Scales

يعرف القياس بأنه الأداة التي يتم بوساطتها التعبير عن الخصائص والسمات بالأرقام . ولأن نوع الاختبار الإحصائي الذي سوف يستخدم لاختبار صحة الفروض، أو أي نوع آخر من أنواع التحليل الإحصائي، يعتمد اعتمادا كليا على المستوى الذي تم فيه قياس المتغير، يكون مهما جدا التعرف على أنواع أو مستويات أو موازين القياس . وهناك أربعة مستويات للقياس نتعرض بالشرح لكل واحد منها فيما يلي :

١.٥.٢.٢ مستوى القياس الاسمي Nominal Measurement Scale

وهو أبسط وأدنى مستويات القياس جميعا حيث يتم بوساطته تسمية أو تصنيف المبحوثين إلى مجموعات متمايزة طبقا لخصائص نوعية (يمكن بالمثل تسميتها كيفية أو اسمية أو وصفية) فيكون الاختلاف بين المبحوثين فقط في الاسم الذي تماثل وظيفته تماما وظيفة الملصق في علبة الدواء، مثلا، لتمييزه عن بقية الأدوية الموضوعة إلى جنبه في الأرفف، ولا تتعدى ذلك قيد أنملة. ومن أمثلة المتغيرات التي يتم قياسها على هذا المستوى من مستويات القياس: متغير الجنس (ذكر، أو أنثى)، ونمط المعيشة (حضر، أو ريف، أو بدو)، والحالة الزواجية (متزوج، أو عزب، أو مطلق، أو أرمل)، وهلم جرا. وإذا ما اختار الباحث إعطاء أرقام لفئات متغير اسمي (أي يقاس على ميزان القياس الاسمي) بغرض تصنيفها: مثلا أعطى الذكور الرقم «١» وأعطي الإناث الرقم «٢»، فإن هذا لا يعني أن «٢» أكبر من «١» أو أن «١» أفضل من «٢». وبمعنى آخر فإن الأرقام هنا ليس لها معنى كمي وإنما تؤدي دور التصنيف (أو الملصق) فقط. أي أن الباحث لا يستطيع إجراء أي عملية حسابية سواء كانت هذه العملية عملية جمع أو طرح أو قسمة أو ضرب على هذه الأرقام ليغير من حقيقة أن هذا ذكر وتلك أنثى. وإنما حدد الرقم «١» ليدل على أن المبحوث «ذكر» وحدد الرقم «٢» ليدل على أن المبحوث «أنثى» - أي حدد ملصقين ليميز بكل واحد منهما الجنس الذي يقصده؛ إذ كان بإمكان الباحث مثلا اختيار الحرفين «أ»، و«ب» للتمييز بين الجنسين بدلا من الرقمين «١» و«٢» ولا يغير ذلك من الأمر شيئا. كما يمكن عند الترتيب أن يجيء الرقم «٢» قبل الرقم «١» أو بعده فالأمر سيان مادام أن القصد هو فقط التمييز بين شيئين وليس السعي إلى تسجيل أفضلية مثلا بين هذين الشيئين لأن مثل هذه الأفضلية منتفية أصلا.

ويجب أن نضع نصب أعيننا ونحن نتعامل مع البيانات أو المتغيرات التي تقاس على ميزان القياس الاسمي ما يلي:

(أ) يجب أن توضع كل حالة (أو مفردة أو مبحوث) من الحالات المدروسة في فئة أو مجموعة واحدة فقط من المجموعات المتمايزة.

(ب) الفئات التي تم الأخذ بها عند التصنيف يجب ألا تكون متداخلة بمعنى ألا نبقي على أي احتمال لإمكانية تصنيف أي حالة في أكثر من مجموعة واحدة نتيجة للبس في تعريف هذه المجموعة أو تلك. مثال ذلك أن يتم تصنيف «أرمل» في فئة «عزب» أو تصنيف «مطلق» في فئة «متزوج»(٥) نتيجة للخلط في الفهم الناتج عن عدم التعريف الدقيق لفئات متغير الحالة الزواجية.

(جـ) يجب أن تكون فئات التصنيف ممثلة وشاملة لجميع المبحوثين (أي لجميع الحالات التي تدرس). وبمعنى آخر فإن كل مفردة أو حالة أو مبحوث يستطيع أن يصنف نفسه تبعا لأحد هذه الفئات ولفئة واحدة فقط. ولسد باب الذرائع أمام التصنيف الملتبس تخلق عادة فئة إضافية تسمى «أخرى» تصنف فيها الحالات التي لا تجد لها متسعا في أي فئة من الفئات المنصوص عليها غير هذه الفئة. وغالبا ما تضم فئة «أخرى» حالات قليلة جدا يتم التعامل معها بطريقة خاصة عند تحليل البيانات.

٢,٢,٥,٢ مستوى القياس الترتيبي Ordinal Measurement Scale

وهو أعلى مرتبة من مستوى القياس الاسمي حيث يمكن باستخدامه ترتيب المفردات أو الحالات المدروسة ترتيبا تصاعديا أو تنازليا حسب درجة امتلاكها لخاصية معينة، بينما تستحيل هذه السمة لدى مستوى القياس الاسمي. ومن أمثلة الأوضاع التي يستخدم فيها مستوى القياس الترتيبي ترتيب مجموعة من الأفراد حسب متغير المستوى المعيشي بفئاته التي تم تحديدها في: مرتفع، متوسط، متدن؛ أو في متغير المستوى التعليمي: أمي، ابتدائي، متوسط، ثانوي، جامعي وتعليم عالٍ؛ أو في متغير مدى درجة الموافقة على مقولة معينة: موافق بشدة، موافق، غير موافق، غير موافق بشدة. ويلاحظ في كل هذه الأمثلة إمكانية إعطاء أرقام للفئات تدريجا من الأصغر إلى الأكبر أو العكس ويكون لهذه الأرقام معنى يتضمن الأفضلية (أي معنى ترتيبيا) إلا أن الفروق أو المسافات بين هذه الأفضليات لا يمكن تحديدها أولا، ولا

---

(٥) يمكن في حالة الإهمال في تدريب جامعي البيانات مثلا أن يفهم أحدهم بأن المطلق «متزوج» بحكم أنه «سبق له الزواج» أو أن الأرمل «عزب» مادام أنه يعيش بدون زوجة.

يمكن الزعم بأنها متساوية ثانيا (أي أنه لا توجد وحدة قياس بهذا المستوى من مستويات القياس). وبصورة أكثر إفصاحا، هب أننا اصطلحنا على الرموز أو الأرقام (٤ ، ٣، ٢ ، ١) لتدل على ترتيب الفئات: موافق بشدة، موافق، غير موافق، غير موافق بشدة. فكلمة موافق بشدة تقابل الرقم (٤) إلا أن هذا لا يمنع من أن تعطى الفئة ذات الرتبة الأعلى، وهي في هذه الحالة «موافق بشدة»، الرقم (١) بدلا من (٤) مع بقاء المعنى الترتيبي واضحا. وهكذا يتضح أن القياس بهذا المستوى لا يكتفي بأن يبين اختلاف الأفراد بالنسبة لسمة أو لخاصية معينة كما هو الحال في مستوى القياس الاسمي، وإنما يذهب إلى أبعد من ذلك فيرتبهم أيضا حسب درجة امتلاكهم لتلك السمة، ولو أن المسافات بين الفئات في هذا المستوى غير متساوية. فمثلا، لا يمكن الزعم بأن المسافة (أو الفرق) بين «موافق بشدة» و«موافق» هي المسافة نفسها بين «موافق» و«غير موافق». فمن الواضح أنها غير متساوية، وذلك لأن في الحالة الأولى يتفق من اختار «موافق بشدة» ومن اختار «موافق» في أن كليهما موافق ولكن درجة الموافقة لدى أحدهما أعلى مما هي لدى الآخر ولكن أن تقرر بكم درجة هي أعلى فهذا مستحيل. أما في الحالة الثانية فإن الاختلاف بين من اختار «موافق» ومن اختار «غير موافق» كبير (أي المسافة بينهما كبيرة) إلى درجة النقيض ولكننا نعجز أيضا عن تحديد كمية هذا الاختلاف (أي نعجز عن تحديد بعد المسافة بين الخيارين). مختصر القول إذن أن هذا المقياس يمتلك خاصية التصنيف التي يمتلكها أيضا المقياس الاسمي بالإضافة إلى خاصية الترتيب التي يفتقدها المقياس الاسمي. إلا أن المقياس الترتيبي لا يمتلك وحدة للقياس.

٣,٥,٢,٢ مستوى القياس الفاصل (أو الفتري) Interval Measurement Scale

القياس بالمستوي الفاصل أرقى من القياس بالمستوى الترتيبي، حيث تحمل الأرقام هنا معنى كميا ويكون الحصول على وحدة القياس بالتالي متاحا. وبعد الاطلاع على المثال الذي يلي سوف نتيقن من أن لهذا المقياس وحدة قياس بالإضافة إلى سمتي التصنيف والترتيب اللتين يتمتع بهما ميزان القياس الترتيبي، وفوق هذا وذاك فإن

نقطة الإسناد «صفر» هي محض افتراض، بمعنى أن «الصفر» هنا لا يعني انعدام الخاصية كما سوف يتضح حالا، وإنما هو «صفر» نسبي وليس مطلقا.

إن علامات الطلاب في مادة ما، مثلا، تمثل متغيرا يقاس على مستوى القياس الفاصل. فإذا كانت علامات الطلاب في مادة الإحصاء مثلا تتوزع بين الصفر والمئة بوحدة الخمس علامات (أي: صفر، ٥، ١٠، ١٥، . . . ، ٩٥، ١٠٠). معنى هذا أن:

(أ) الطلاب يختلفون في تحصيلهم، فهذا حصل على ٢٥ علامة وذاك حصل على ٥٥ وثالث حصل على ٨٥، وهكذا. وهذا الاختلاف يمكن أن يوفره المستوى الاسمي.

(ب) رتبة الطالب الذي علامته ٧٠ أعلى من رتبة الطالب الذي علامته ٦٥ مثلا، وهذا يمكن أن يوفره المستوى الترتيبي.

(جـ) الطالب الذي علامته ٧٠ أعلى تحصيلا بعشر (١٠) نقاط (أي: وحدتين) من الطالب الذي علامته ٦٠، أي أن الفواصل بين الفئات متساوية أو أن المسافات بين المجموعات المتمايزة (أو الفئات) متساوية، وهذا ما يوفره المستوى الفاصل أو الفتري (وسمي الفتري أيضا لأن هناك فترات أو مسافات بين التصنيفات).

(د) إذا حصل طالب ما على علامة «صفر» في هذه المادة فهذا لا يعني أن الطالب لا يفقه شيئا مطلقا في هذه المادة وإنما لكون ما يعرفه فيها لم يرد في ورقة الاختبار. ولذلك فإن «الصفر» الذي تحصل عليه هو صفر نسبي وليس مطلقا، وهذا ما يوفره المقياس الفتري كذلك.

٤.٥.٢.٢ مستوى القياس النسبي Ratio Measurement Scale

وهو أرقى من مستويات القياس الثلاثة السابقة جميعها حيث يتفوق على مستوى القياس الفتري بأنه يمتلك سمة «الصفر» المطلق الذي يدل على انعدام الخاصية أو السمة، ويتفوق على المستويين الباقيين باشتماله على جميع سماتها بالإضافة إلى سماته الخاصة به. ومثال المتغيرات التي تُقاس على هذا المستوى من القياس: عدد الأبناء بالأسرة، عدد التلاميذ بالصف، عدد الكراسي بحجرة الدراسة. . . إلخ. ففي المثال الأول يمكن أن يكون عدد الأبناء بالأسرة «صفرا» وهو صفر مطلق كما هو

واضح ، وأن الأسرة التي لديها ستة أبناء على سبيل المثال يكون بحوزتها ضعف عدد الأبناء الذين لدى أسرة بها ثلاثة أبناء فقط . ويمكنك اشتقاق ما شئت من الأوضاع لتبين بعض أو كل خصائص مستويات القياس الأخرى المتضمنة في هذا المستوى من مستويات القياس .

٢,٢,٥,٥ خاتمــــة

لقد ذكرنا في بداية الحديث عن مستويات القياس أن الهدف من تقسيمها بهذه الكيفية هو إتاحة إمكانية اختيار الأسلوب المناسب من أساليب التحليل الإحصائي التي تتواءم وكل مستوى من مستويات قياس مختلف أنواع البيانات . ونعيد التأكيد هنا على أن هذا التقسيم لمستوى قياس المتغيرات لا غنى عنه البتة عند التعامل مع مختلف أنواع البيانات بغرض استخراج المؤشرات الإحصائية المناسبة لوصفها أو لاختيار معاملات اختبارات الفروض التي تتناسب وكل نوع من أنواع هذه البيانات . ومفصل القول يتلخص في أن على هذا التقسيم تتوقف معظم العمليات الإحصائية . فمثلا ، لا يمكن حساب المتوسط الحسابي كمقياس من مقاييس النزعة المركزية من بيانات تم قياسها على مستوى القياس الاسمي بينما يمكن ذلك في حالة البيانات التي قيست على المستوى الفتري أو النسبي . كما أن معاملات الارتباط (انظر الفصل السادس) التي تلائم المستوى الفاصل (أي : الفتري) تختلف عن تلك التي تستخدم في حالة المستوى الاسمي أو الترتيبي . فبينما يمكن استخدام معامل ارتباط بيرسون (انظر الفصل السادس) لقياس قوة واتجاه العلاقة بين المتغيرات المقاسة على المستوى الفاصل لا يمكن استخدام النوع نفسه من معاملات الارتباط لوصف العلاقة بين المتغيرات المقاسة على المستوى الاسمي ، بل لا بد من استخدام أحد المعاملات اللامعلمية nonparametric مثل «مربع كاي» أو «جاما» وغيرها (انظر الفصل السابع) . ولذا فإن معرفة مستويات القياس عملية أساسية ومهمة جدا للباحث ، إذ بدونها تصبح كافة العمليات والجهود الإحصائية مجرد مضيعة للوقت والجهد .

وأخيرا يجب الإشارة إلى أن المقياس الجيد لا بد وأن تتوافر فيه صفتان رئيستان هما :

(أ) الصدق Validity .

(ب) الثبات Reliability .

أما «الصدق» فيعني أن المقياس الذي يستخدمه الباحث يقيس بالفعل ما ينبغي أن يقاس . فالمقياس الصادق هو المقياس الذي يقيس ما أعد لقياسه ، أو الذي يحقق الغرض الذي أعد لأجله . وفيما يتعلق بـ«الثبات» فإن المقياس الثابت هو المقياس الذي يقيس بدرجة مقبولة من الدقة precision – أي أنه باستخدام المقياس نفسه والمبحوثين أنفسهم يتم الحصول على النتائج نفسها من وقت لآخر .

٢,٣ أسئلــة

١- كيف تعرّف عملية جمع البيانات؟

٢- أتيح لك بنك معلومات ما من خلال شبكة «الإنترنت» فأخذت منه بيانات حول قضية تهمك . ما اسم نوع البيانات التي أخذتها من حيث مصدر حصولك عليها؟

٣- باحث اجتماعي أراد دراسة إحدى مظاهر تحول اجتماعي بدولة «موزمبيق» بالجنوب الإفريقي . أي وسيلة من وسائل جمع البيانات ميدانيا تنصحه باستخدامها؟ ولماذا؟

٤- تحدث بإيجاز عن «المناخ الصحي» لـ«المقابلة الشخصية» والمواصفات المطلوب توافرها في أداة جمع البيانات المصاحبة .

٥- عدّد مزايا الحصر بالعينة مشيرا إلى الحالات التي لا يمكن للحصر الشامل أن يكون فيها بديلا له .

٦- ما هي المحكات التي يتوقف عليها اختيار أسلوب معين من بين أسلوبي جمع البيانات ميدانيا من حيث درجة شمول هذه البيانات لمفردات مجتمع الدراسة؟

٧- لِمَ سميت العينة الاحتمالية بهذا الاسم بالذات؟ وما أهم ما يميزها عن العينة غير الاحتمالية؟

٨- متى بالضبط نطمئن إلى تعميم نتائج العينة على المجتمع كله؟

٩- أذكر ثلاثة من أنواع العينات غير الاحتمالية وعرّف كل واحدة منها في إيجاز .

١٠- ما هو الفرق بين «المجتمع الإحصائي» و«المجتمع الإنساني»؟
١١- تحدث بإيجاز عن مستويات القياس وبين لماذا يجب أن يكون لدينا مستويات لقياس المتغيرات أصلا.

٢,٤ اصطلاحات ينبغي تذكرها

- مصادر جمع المعلومات.
- الاستبانة.
- المقابلة الشخصية.
- الحصر الشامل/ الحصر بالعينة.
- المجتمع الإحصائي.
- العينة الاحتمالية/ العينة غير الاحتمالية.
- المتغير الكمي/ المتغير النوعي.
- المتغير المتصل/ المتغير المنفصل.
- موازين القياس.
- الصدق.
- الثبات.

الفصل الثالث

عرض البيانات
Presentation of Data

٣,١ مقدمـة

إن البيانات التي يتم جمعها من المبحوثين مباشرة سواء كانت في صورة عبارات أو أرقام لا تفضح عن أي معنى ذي بال لمن يحاول فهمها في صورتها المبدئية تلك، ولهذا أطلق عليها مسمى «بيانات خام» raw data . ولكي ندلل على صحة ما ذهبنا إليه من تعذر فهم ما تنبض به الأرقام «الخام» أو العبارات «الخام» دعنا نتأمل ذلك، على الترتيب، من المثالين (١ , ٣)، (٢ , ٣).

المثال (٣,١). علامات خمسين طالبا في مادة الإحصاء الاجتماعي.

٧٠	٥٥	٥١	٤٢	٥٧
٤٥	٦٠	٤٧	٦٣	٥٣
٣٣	٦٥	٣٩	٨٢	٥٥
٦٤	٨٨	٦١	٦٥	٤٢
٥٠	٥٢	٥٣	٤٥	٥٥
٢٥	٣٦	٥٩	٦٣	٣٩
٦٥	٤٥	٤٩	٢٠	٦٤
٧٥	٤٢	٤١	٥٢	٧٨
٣٠	٣٥	٢٥	٤٨	٢٦
٥٤	٤٠	٥٥	٤٦	٥٨

المثال (٣,٢). الحالة الزواجية لعشرين موظفا.

متزوج	عزب	متزوج	متزوج	عزب
مطلق	أرمل	عزب	مطلق	أرمل
متزوج	متزوج	عزب	أرمل	متزوج
مطلق	متزوج	مطلق	عزب	أرمل

إذا علمت أن الدرجة الكاملة لاختبار مادة الإحصاء الاجتماعي تتألف من مئة نقطة، هل بمستطاعك التعليق على المستوى العام لهؤلاء الطلاب في هذه المادة بنظرة سريعة على علاماتهم المرصودة بالمثال (١ ,٣)؟ هل بإمكانك أن ترصد بسهولة الطلبة ذوي الأداء الضعيف الذين تقل علاماتهم عن الخمسين؟ وبالمقابل هل بمستطاعك أن تعيّن للتو الطلاب جيدي الأداء الذين تفوق علاماتهم السبعين؟ إن الإجابة على جميع هذه الأسئلة يرجح فيها النفي، إذ إنه يصعب على المرء أن يقيم أداء الطلاب في هذا الاختبار بمسح سريع للعلامات المدونة بهذه الصورة الخام كما هو الحال في المثال (١ ,٣). ويتطلب الأمر بعضا من الوقت للقيام بعملية حصر مملة تجرى على أرقام لا يبدو أنها توفر عنصر إمتاع بين ثناياها. ولسوف تتبيّن في مقبل طرحنا هذا أنه بإمكانك الاستعانة بأشكال هندسية ورسوم بيانية وجداول تنظيمية تجعل من مثل هذه البيانات الخام مادة جدّ مشوقة تشدك إليها، بالإضافة إلى أنها سوف تمكنك من الإجابة على الأسئلة السابقة بمجرد النظر إلى هذه الأشكال الهندسية والرسوم البيانية والجداول التنظيمية.

إن ما ينطبق على المثال (١ ,٣) من صعوبة في التقييم الأولي للظاهرة بمجرد النظر إلى المادة الخام ينطبق أيضا على المثال (٢ ,٣) رغم أن الأخير يتضمن حالات قليلة نسبيا مما يجعل المهمة أسهل نسبيا إلا أنها قطعا أعقد بكثير مما لو وضعت بيانات المثال (٢ ,٣) في صورة جدول تنظيمي مثلا. والحال أنه ليس سهلا الحكم فيما إذا كانت ظاهرة الطلاق تكتنف هؤلاء الموظفين أكثر أو أقل مما تكتنفهم ظاهرة الترمل بنظرة عابرة على المثال (٢ ,٣)، أو أن المتزوجين أقل أو أكثر من العُزّاب، مثلا، بمجرد النظر إلى هذه البيانات الخام في المثال نفسه. ونظرا لما سقناه من دليل على صعوبة فهم البيانات الخام ولأنه من النادر استخلاص أي نتائج مفيدة من مثل هذه البيانات في صورتها الابتدائية تلك، كانت الحاجة ماسة إلى وظيفة للإحصاء تتلافى هذا القصور. وهكذا برزت وظيفة عرض البيانات التي تتيح سبلا تتنافس في إعطاء صورة واضحة للبيانات يسهل على القارئ فهمهما دون عناء. إذًا، فإن تنظيم وعرض البيانات من الخطوات الأساسية التي تفضي إلى سهولة تدبر ما تفصح عنه فتهيئ بذلك للخطوات اللاحقة من تحليل واستنتاج. وتستخدم عدة طرق لعرض البيانات تتمثل في عرضها في رسوم بيانية تشمل أشكالا هندسية رباعية أو مثلثة وصور تعبيرية ومضلعات ومنحنيات، كما يتم

عرضها أيضا في صورة جداول تنظيمية وتلخيصية تتضمن أحجام الظاهرة بجوانبها المختلفة وتكراراتها النسبية والنسبية المئوية وكذلك تكراراتها التجمعية . وسوف نعرض إلى ذلك كله تدريجيا من البيانات (أو الإحصاءات أو المعلومات أو القراءات) التي يتضمنها جدول إحصائي بسيط إلى الإحصاءات الأكثر تعقيدا في عرضها الجدولي .

٣,٢ العرض البياني للقراءات Data Graphic Presentation

سوف نستخدم القراءات التي يتضمنها الجدول رقم (١ , ٣) لاستعراض أهم طرق العرض البياني . وكما هو واضح ، فإن هذا الجدول الإحصائي البسيط يبين عدد الطلاب والطالبات المسجلين في إحدى المدارس الأهلية بمدينة ما في الفترة من ١٤١٠هـ إلى ١٤١٥هـ .

الجدول رقم (٣,١). عدد الطلاب والطالبات المسجلين في إحدى المدارس الأهلية.

العام لدراسي	١٤١٠/١٤١١هـ	١٤١١/١٤١٢هـ	١٤١٢/١٤١٣هـ	١٤١٣/١٤١٤هـ	١٤١٤/١٤١٥هـ
الطلاب	٦٠٠	٨٠٠	١٠٠٠	١٢٠٠	١٤٠٠
الطالبات	١٥٠	٢٥٠	٤٠٠	٦٠٠	٩٠٠
المجموع	٧٥٠	١٠٥٠	١٤٠٠	١٨٠٠	٢٣٠٠

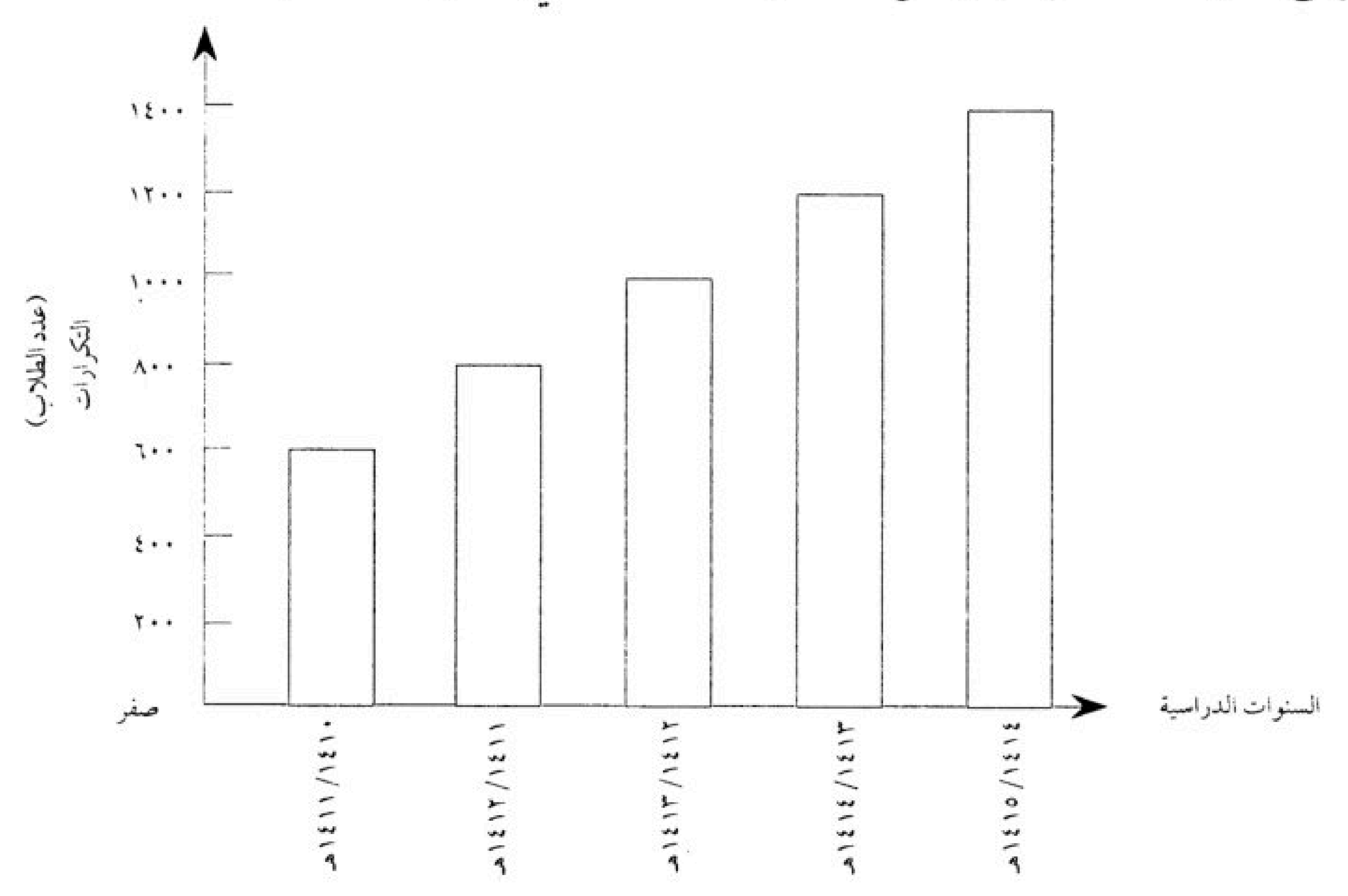
فإذا طلب منا ما يلي :

١ – عرض مجموعات الطلاب باستخدام الأعمدة البسيطة .
٢ – عرض مجموعات الطلاب والطالبات باستخدام الأعمدة المزدوجة .
٣ – عرض مجموعات الطلاب والطالبات باستخدام الأعمدة المجزأة .
٤ – عرض مجموعات الطلاب باستخدام الخط البياني .
٥ – عرض مجموعات الطلاب والطالبات باستخدام خريطة الشريط .
٦ – عرض مجموعات الطلاب باستخدام قطاعات الدائرة .

نشرع في الحل فنبدأ بالأعمدة البسيطة .

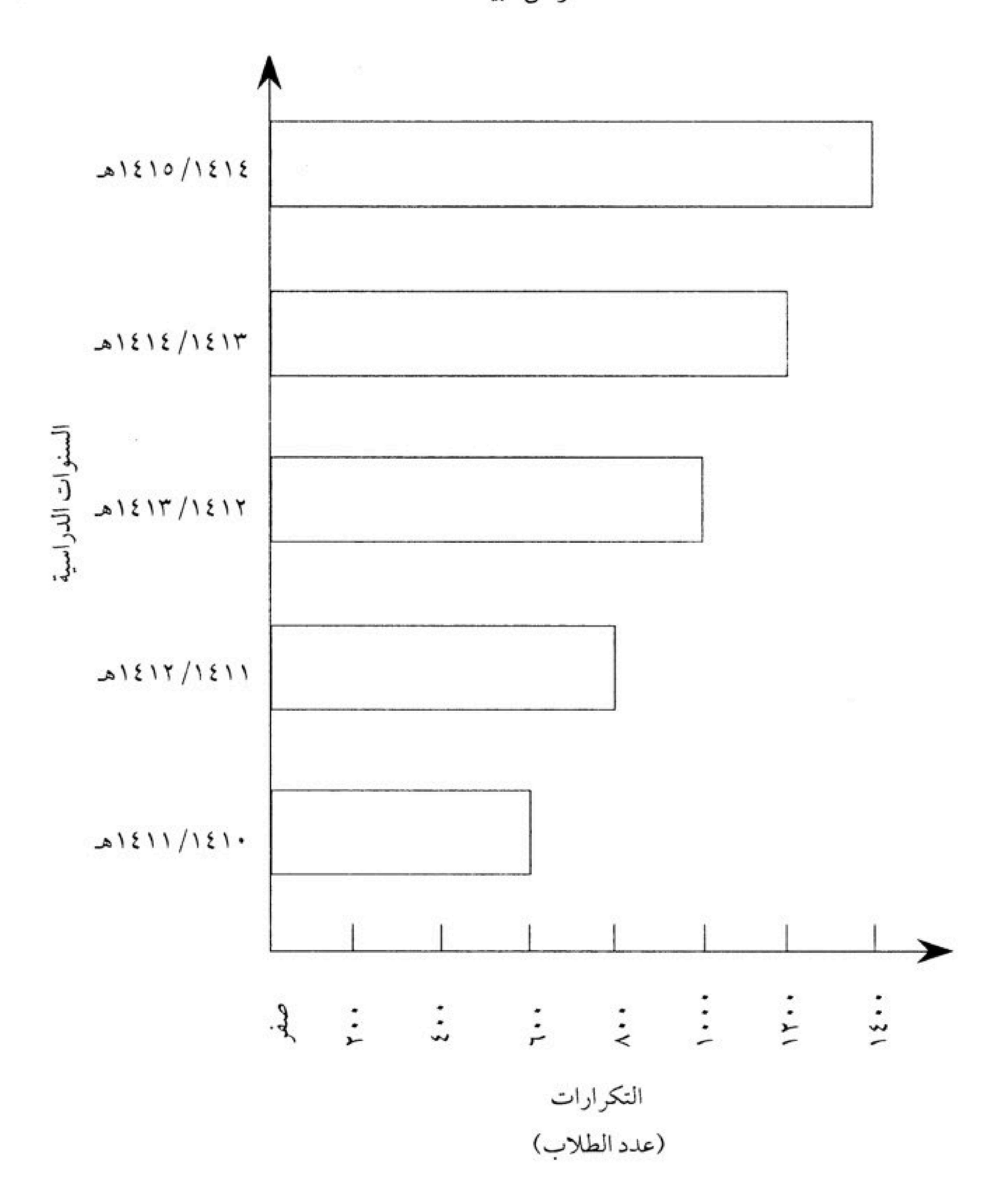
١,٢,٣ الأعمدة البسيطة Bar Graphs

لعرض قراءات الظاهرة باستخدام الأعمدة البسيطة نبدأ برسم محورين أحدهما أفقي والآخر رأسي. يمثل المحور الأفقي أوجه أو نقاط الظاهرة المختلفة (السنوات الدراسية في مثالنا هذا) ويمثل المحور الرأسي على كل نقطة من نقاط الظاهرة شكلا رباعيا (عمودا) يتناسب ارتفاعه مع عدد التكرارات المقابلة لكل نقطة على أن تكون للأشكال الرباعية الناتجة قواعد بأطوال متساوية حسب مقياس الرسم المختار كما يتعين أن تكون المسافات بين مختلف الأشكال الرباعية أيضا متساوية. الشكل رقم (١ , ٣أ) يتولى استيفاء المطلب (١) من المطالب المتسلسلة في الصفحة السابقة.



الشكل رقم (١.٣أ). توزيع الطلاب المسجلين في الفترة (١٤١٠–١٤١٥هـ).

وهكذا كما نرى في الشكل رقم (١ , ٣أ) فإن الأعمدة البسيطة تستخدم لتمثل ظاهرة واحدة فقط. ويمكن بالطبع أن ترسم الأشكال الرباعية على المحور الرأسي الذي سيشكل قاعدة لها في هذه الحال وتوضع التكرارات على المحور الأفقي حيث يتبادل المحوران مواقفهما التي في الشكل رقم (١ , ٣أ) ليبدو الرسم الجديد كما هو في الشكل رقم (١ , ٣ب).

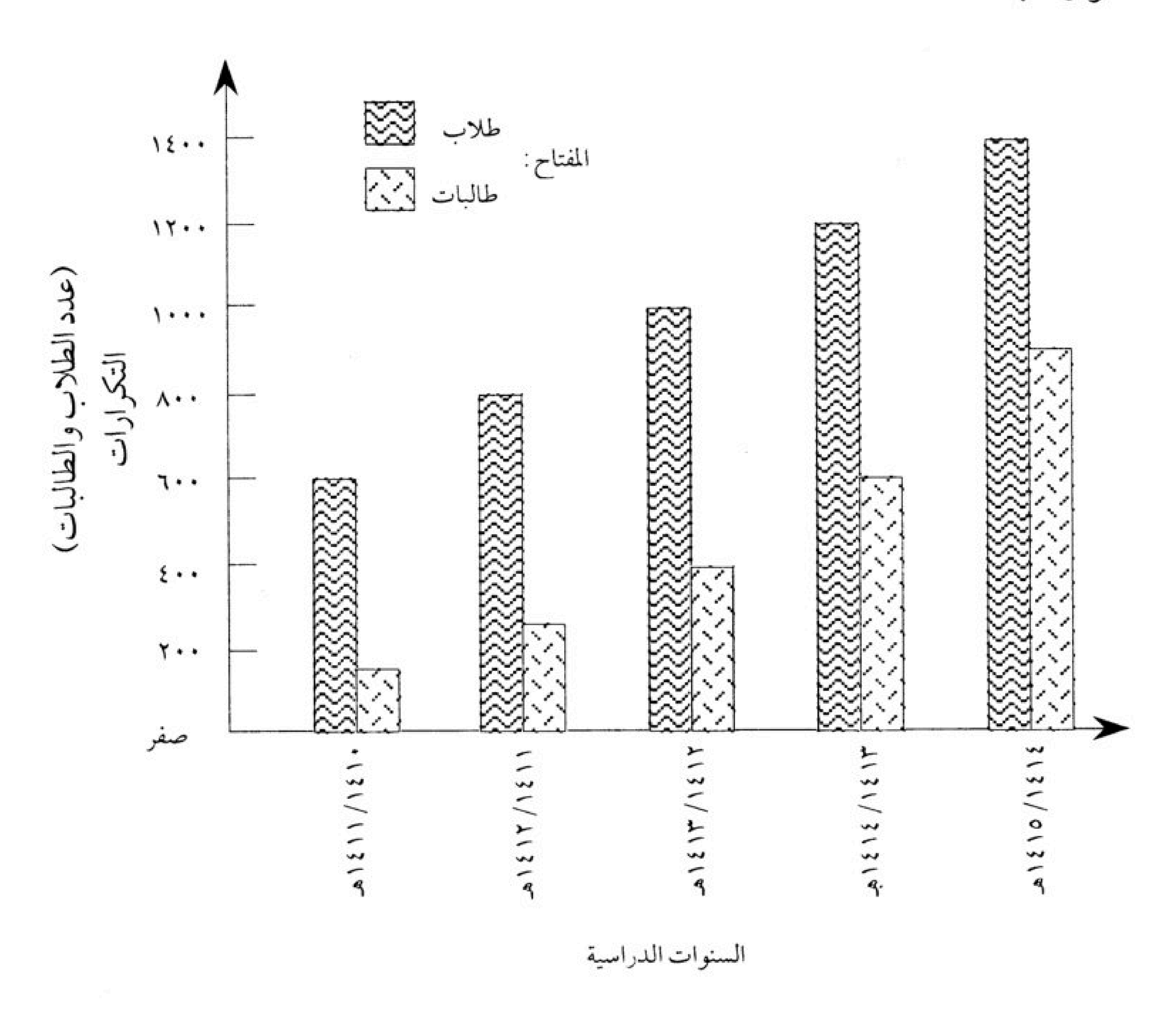


الشكل رقم (١,٣ب). توزيع الطلاب المسجلين في الفترة (١٤١٠-١٤١٥هـ).

٣,٢,٢ الأعمدة المزدوجة Pair Bar Graphs

وتستخدم الأعمدة المزدوجة لعرض ظاهرتين مشتركتين في خاصية أو أكثر كخاصيتي الزمان والمكان في مثالنا هذا أو أي وجه آخر. وكل زوج من القيم المتناظرة للظاهرتين (مجموعتي الطلاب والطالبات في مثالنا هذا) يمثل بوساطة شكلين رباعيين متجاورين. ولتمييز إحدى الظاهرتين عن الأخرى نستخدم لكل واحد منهما لونا أو

ظلا مختلفا . والشكل رقم (٢ , ٣) يوضح كيفية استخدام الأعمدة المزدوجة لعرض ظاهرتين مقترنتين بوجه إحصائي ما كمثال نسوقه استجابة للمطلب (٢) من المطالب الواردة بالمسألة .

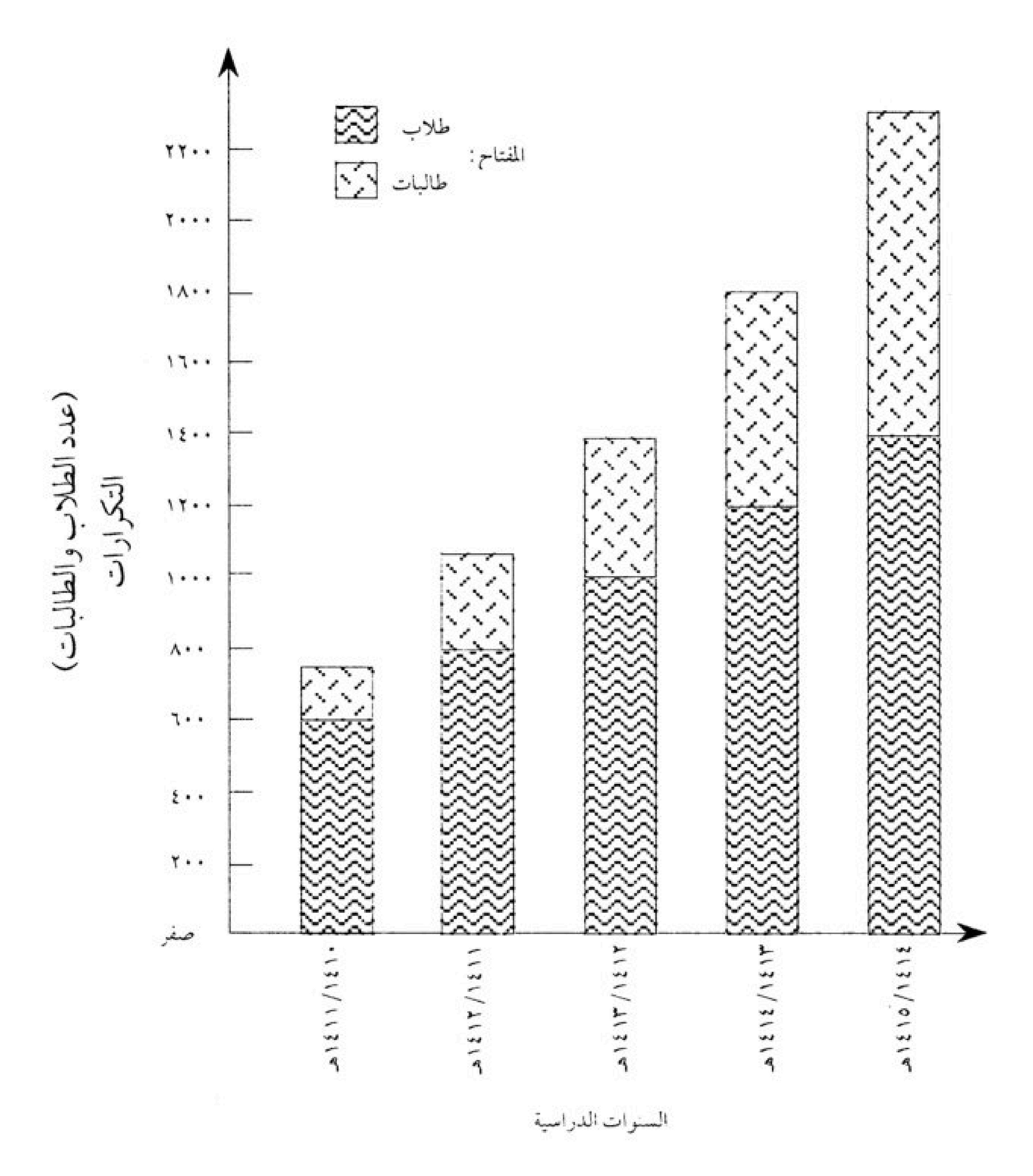


الشكل رقم (٣,٢). توزيع الطلاب والطالبات الذين تم تسجيلهم في الفترة (١٤١٠-١٤١٥هـ).

٣,٢,٣ الأعمدة المجزأة Superimposed Bar Graphs

وتستخدم الأعمدة المجزأة كذلك لعرض ظاهرتين تشتركان في خاصية أو أكثر إلا أن الظاهرتين هنا تمثلان في شكل رباعي (عمود) واحد تتم تجزئته إلى قسمين ويعطى كل واحد منهما لونا أو تظليلا يختلف عن القسم الآخر . ويوضح الشكل رقم (٣ , ٣) الرسم البياني الذي يمثل هذا النوع من الأعمدة المجزأة بعرض توزيع الطلاب والطالبات

على السنوات الدراسية، والذي يعرض استجابة للمطلب (٣) الوارد بالمسألة المعنية.

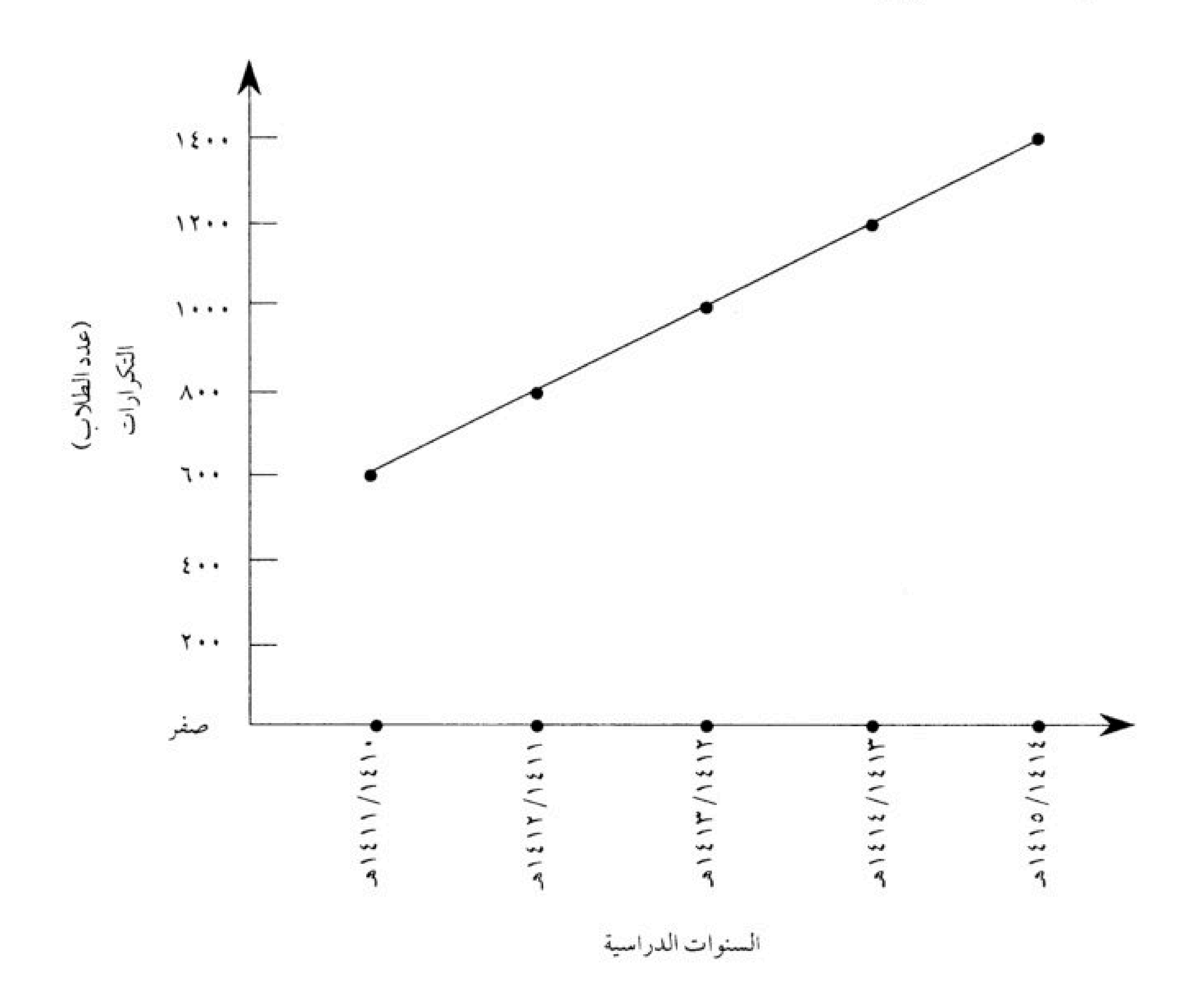


الشكل رقم (٣,٣). توزيع الطلاب والطالبات في الفترة (١٤١٠-١٤١٥هـ).

٣,٢,٤ عرض ظاهرة باستخدام «الخط البياني» Polygon

تمثيل ظاهرة واحدة بوساطة الخط البياني هو عمل مشابه لتمثيلها باستخدام الأعمدة البسيطة (انظر الشكل رقم ١ , ٣) إلا أن الأعمدة هنا يستعاض عنها بنقاط تشير إلى قمم هذه الأعمدة، والتي تتناسب وتكرارات هذه الظاهرة، ثم يصار إلى وصل هذه النقاط مع بعضها فنتحصل بذلك على خط يبين اتجاه الظاهرة صعودا أو هبوطا مع

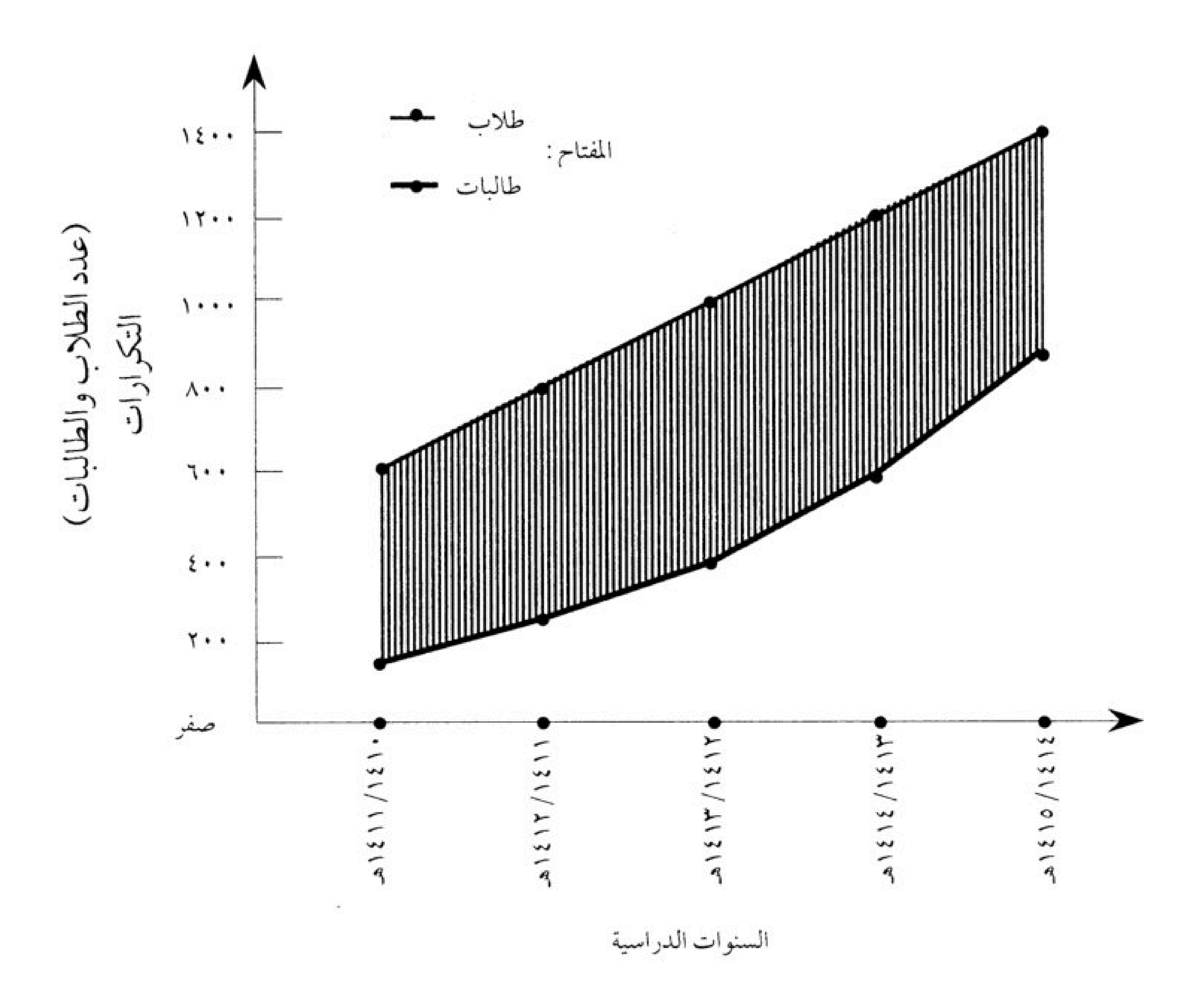
خاصية معينة بمرور الزمن (في مثالنا عدد الطلاب السنوي تبعا للأعوام الدراسية). الشكل رقم (٤ , ٣) رُسِمَ استجابة للمطلب (٤) الوارد بالمسألة.



الشكل رقم (٣,٤). توزيع الطلاب في الفترة (١٤١٠-١٤١٥هـ) باستخدام الخط البياني.

٣,٢,٥ عرض ظاهرتين باستخدام خريطة الشريط Strip Mapping

التمثيل البياني لظاهرتين باستخدام خريطة الشريط يتصف بنهج مشابه لنهج الخط البياني، إذ يرسم خط بياني يمثل إحدى الظاهرتين وآخر يمثل الظاهرة الأخرى على أن يتم التمييز بينهما بتنقيط مسار أحدهما مثلا وإبقاء الآخر متصلا عند وصل النقاط البيانية لكل ظاهرة من الظاهرتين. ثم يصار إلى تظليل المساحة المحصورة بين الخطين لتمثل بذلك خريطة الشريط. والشكل رقم (٥ , ٣) الذي يلي يعطي مثالا للتمثيل البياني باستخدام خريطة الشريط، وهو يلبي المطلب (٥) من قائمة المطالب بالمسألة الخاصة بالطالبات والطلاب المسجلين بالمدرسة الأهلية للأعوام المرصودة.



الشكل رقم (٣,٥). توزيع الطلاب والطالبات في الفترة (١٤١٠-١٤١٥هـ) باستخدام خريطة الشريط.

٣,٢,٦ عرض ظاهرة باستخدام الدائرة البيانية Pie Chart

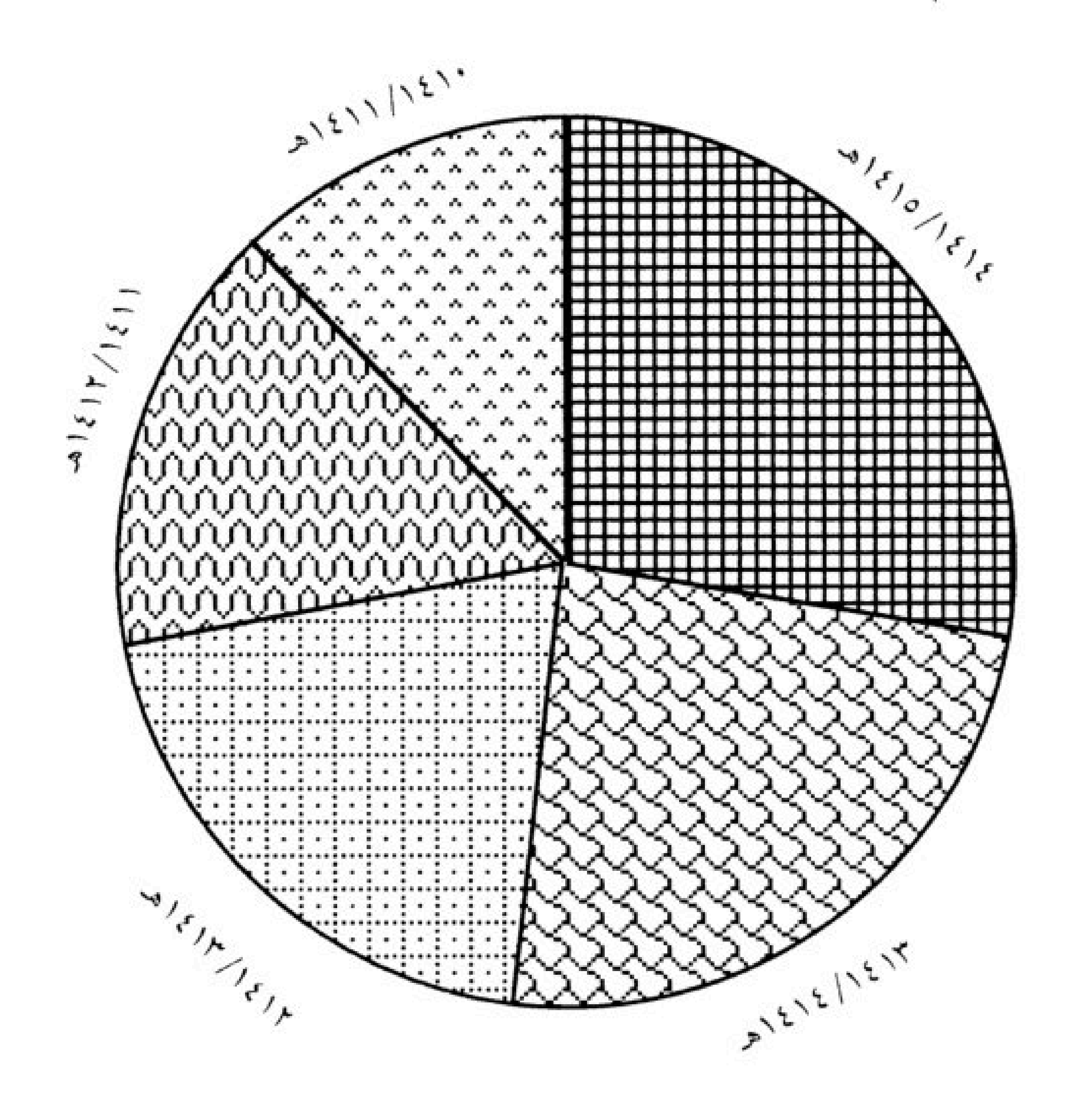
يمكن استخدام الرسوم الدائرية في عرض ظاهرة واحدة إذا كان مجموع مفردات هذه الظاهرة مقسم إلى مجموعات متمايزة تمثل أجزاء فرعية لها. وإذ تمثل المساحة الكلية للدائرة مجموع تكرارات الظاهرة، تقسم هذه المساحة إلى قطاعات تتلاقى رؤوسها في مركز هذه الدائرة بحيث تكون مساحات هذه القطاعات متناسبة مع مجاميع التكرارات الجزئية أو الفرعية التي تشكل في مجموعها المجموع الكلي لمفردات الظاهرة. ويتم تمييز هذه القطاعات بعضها عن بعضها الآخر بتظليلها أو بتلوينها بألوان مختلفة. وحيث أننا نعلم أن الزاوية المركزية للدائرة تساوي ٣٦٠ درجة (٣٦٠°) نقوم بتقسيم مجموع هذه الدرجات إلى مئة (١٠٠) جزء متساوٍ بحيث نحصل على معامل ضرب

بقيمة ٦ , ٣ (أي ٣٦٠ ÷ ١٠٠ = ٦ , ٣). معامل الضرب هذا هو الذي نستخدمه في الحصول على مقدار زاوية كل قطاع من القطاعات التي تمثل التكرارات الجزئية، وذلك بضربه في النسبة المئوية لكل تكرار جزئي من مجموع تكرارات الظاهرة. وللحصول على هذه النسب المئوية الجزئية؛ أولا نقوم بإيجاد قيمة تناسب كل تكرار جزئي إلى المجموع الكلي للتكرارات ثم نقوم بضرب كل قيمة من قيم التناسب الجزئية في مئة لنتحصل على النسبة المئوية التي تضرب بدورها في معامل الضرب (٦ , ٣) لنتحصل أخيرا على زاوية كل قطاع من القطاعات التي تمثل الظاهرة في مجملها. ولنستعين الآن بتكوين جدول يدعم عمليا هذا الطرح النظري ونحن نقوم بتمثيل الطلاب في مثالنا بوساطة الرسم الدائري استجابة للمطلب السادس والأخير بالمسألة (انظر الجدول رقم «٢ , ٣»).

جدول رقم (٣.٢). حساب الزوايا القطاعية للطلاب.

العام الدراسي	عدد الطلاب	قيمة التناسب	النسبة المئوية (٪)	الزاوية بالدرجات
١٤١٠/١٤١١هـ	٦٠٠	٦٠٠ / ٥٠٠٠	٦٠٠ / ٥٠٠٠ × ١٠٠ = ١٢	١٢ × ٣,٦° = ٤٣,٢°
١٤١١/١٤١٢هـ	٨٠٠	٨٠٠ / ٥٠٠٠	٨٠٠ / ٥٠٠٠ × ١٠٠ = ١٦	١٦ × ٣,٦° = ٥٧,٦°
١٤١٢/١٤١٣هـ	١٠٠٠	١٠٠٠ / ٥٠٠٠	١٠٠٠ / ٥٠٠٠ × ١٠٠ = ٢٠	٢٠ × ٣,٦° = ٧٢°
١٤١٣/١٤١٤هـ	١٢٠٠	١٢٠٠ / ٥٠٠٠	١٢٠٠ / ٥٠٠٠ × ١٠٠ = ٢٤	٢٤ × ٣,٦° = ٨٦,٤°
١٤١٤/١٤١٥هـ	١٤٠٠	١٤٠٠ / ٥٠٠٠	١٤٠٠ / ٥٠٠٠ × ١٠٠ = ٢٨	٢٨ × ٣,٦° = ١٠٠,٨°
المجموع	٥٠٠٠	-	١٠٠	٣٦٠°

وبقراءة زاوية كل قطاع من القطاعات الخمس كما هي محسوبة بالجدول رقم (٢ , ٣) نقدم الآن على رسم الدائرة البيانية كما هي مبينة على الرسم البياني الذي يعرضه الشكل رقم (٦ , ٣).



الشكل رقم (٣,٦). توزيع الطلاب في الفترة (١٤١٠-١٤١٥هـ) باستخدام الدائرة البيانية.

٣,٣ العرض الجدولي للبيانات Tabulation

يضطلع العرض الجدولي للبيانات بمهمة تصنيف وتبويب - أي تنظيم - هذه البيانات بصورة تجعلها أقرب لفهم الشخص العادي، كما يمهد هذا العرض لتقديم البيانات بكيفية أكثر إيضاحا باستخدام الرسومات البيانية علاوة على التمهيد لحساب المؤشرات الإحصائية مثل النسب والمعدلات والنسب المئوية ومقاييس النزعة المركزية ومقاييس التشتت. ولتيسير فهم أعمق لعملية عرض البيانات جدوليا سوف نقوم باستعراض موجز لأنواع الجداول الإحصائية من بسيطة ومركبة أو مزدوجة، ثم نستعرض بعدئذ تصنيف البيانات إلى بيانات وصفية (أو نوعية أو كيفية أو اسمية) وإلى بيانات كمية قبل الشروع

في تناول عملية الجدولة التي تأخذ نمطا مختلفا إلى حد ما تبعا لاختلاف نوع البيانات التي نحن بصدد جدولتها. وحتى البيانات الكمية تتطلب تقسيمها إلى بيانات كمية منفصلة discrete وبيانات كمية متصلة continuous، ثم إلى بيانات كمية ذات مدى صغير نسبيا وأخرى ذات مدى كبير نسبيا؛ وذلك للتمكن من تأطير عملية الجدولة في إطار وضعها الأمثل.

١.٣.٣ أنواع الجداول الإحصائية

يتضمن العرض الجدولي نوعين من أنواع الجداول الإحصائية أيًا يكن نوع البيانات الذي يعرض. أحد هذين النوعين يسمى الجدول الإحصائي البسيط simple table، والنوع الآخر يسمى الجدول الإحصائي المركب أو المزدوج compound table. ويتمثل الفرق بينهما في أن الجدول الإحصائي البسيط يتضمن مجموعة من الحالات المدروسة يتم توزيعها طبقا لفئات متغير واحد فقط، بينما يتضمن الجدول الإحصائي المزدوج مجموعة من الحالات المدروسة يتم توزيعها طبقا لفئات متغيرين أو أكثر. وتأسيسا على هذا الفرق، فإن شكل الجدول الإحصائي البسيط يبدو في عمودين فقط: أحدهما يعكس أوجه الظاهرة على اختلافها (أي يعكس مسميات فئات المتغير المعني) والآخر يعرض نواحيها الكمية كما سوف يتضح جليا في أمثلة لهذا النوع من الجداول نوردها بعد قليل. أما الجدول الإحصائي المركب فيحتل أكثر من عمودين مادام أن مجموعة الحالات المدروسة يجري تقسيمها وفقا لأوجه ظاهرتين أو أكثر (أي تبعا لفئات متغيرين أو أكثر) كما سوف يتضح أيضا من الأمثلة اللاحقة.

وكثيرا ما تجد في كتب الإحصاء المنهجية الإشارة إلى الجدول الإحصائي بأنه جدول تكراري frequency table وما ذلك إلا لأن الجدول الإحصائي يعرض الظاهرة (أي المتغير) في صورة تكرارات أو مجموعات، كل مجموعة أو تكرار يقابل وجها معينا (أي فئة معينة) من أوجه تلك الظاهرة (أي يقابل فئة معينة من فئات المتغير الذي بموجب فئاته تم تقسيم المجموع الكلي للتكرارات) مثل تقسيم أو تصنيف عدد من الأشخاص إلى ذكور وإناث تبعا لمتغير الجنس. وتكمن أهمية الجدول التكراري في أنه يتمتع بمزايا يمكن إيجازها فيما يلي:

١- تلخيص البيانات، حيث يتم عرضها في جدول صغير مهما كان عدد البيانات أو القراءات المتحصل عليها.

٢- يؤدي هذا التلخيص إلى إفصاح عن المعلومات بصورة مباشرة وسريعة ويساعد ترتيب البيانات على إمكانية ذلك الإفصاح الذي - لولا الجدول التكراري - لم يكن ممكنا بالنظر إلى أعداد كبيرة من القيم المتناثرة والمتباعدة وغير المرتبة (انظر المثال «١ , ٣» السابق).
٣- إمكانية المقارنة بين مجموعتين أو أكثر بعرضها في جدول واحد.
٤- يمكن حساب كافة المقاييس الإحصائية (مقاييس النزعة المركزية والتشتت) من هذا الجدول المختصر بدلا من الرجوع إلى البيانات الأصلية الكبيرة العدد، وفي ذلك تسهيل كبير لحساب هذه المقاييس.
٥- بعض المقاييس الإحصائية يلزم لحسابها أن توضع البيانات في جدول تكراري.
٦- إمكانية عرض الظاهرة محل البحث عرضا بيانيا.

٣,٣,٢ أسس تصنيف البيانات عند الجدولة

٣,٣,٢,١ التصنيف النوعي أو الكيفي

نلجأ إلى التصنيف حسب الصفات المميزة لكل مجموعة (تكرار) من مجموعات الحالات المدروسة إذا كانت الخاصية التي تشترك فيها المفردات (أي الحالات المدروسة) غير قابلة للقياس. أمثلة ذلك توزيع المفردات حسب نوع الشهادة (انظر الجدول البسيط رقم ٣ , ٣) أو حسب الحالة الزواجية، أو الجنس، أو حسب أي ظاهرة أخرى لا تخضع للقياس.

٣,٣,٢,٢ التصنيف الزمني (التأريخي)

يتم الرجوع إلى هذا التصنيف عندما تتكون البيانات التأريخية (أو السلاسل الزمنية) من أرقام تتعلق بالمقادير التي تأخذها ظاهرة معينة في أوقات متتابعة مثل مجموعات الحجاج القادمين إلى السعودية في سنوات مختارة (انظر الجدول التكراري البسيط رقم ٤ , ٣)، أو كمية الإنتاج من محصول معين لدولة ما في سنوات متتالية، أو عدد السجناء في مدينة ما في سنوات متتالية، أو حوادث الحركة في دولة ما لأسابيع متتالية، وهلم جرا.

٣,٢,٣,٣ التصنيف الكمي

عندما تشترك المفردات (أو الحالات المدروسة) في خاصية يمكن قياسها أو عدّها يصبح المقدار أو الكم quantity هو الأساس الذي يتم بموجبه التصنيف، مثل الطول، والوزن، والعمر، ودرجة الحرارة . . . إلخ. وكأمثلة ينطبق عليها هذا النوع من التصنيف، انظر أيضا الجداول التكرارية البسيطة أرقام (٥ ,٣)، (٦ ,٣) و(٧ ,٣).

٣,٣,٢,٤ التصنيف الجغرافي

تجد في كثير من الأحيان اعتبار التصنيف الجغرافي تصنيفا مستقلا مع وجود ما يسوغ اعتباره في حقيقة الأمر فرعا من التصنيف النوعي qualitative السابق ذكره. وكمثال للتصنيف الجغرافي انظر الجدول التكراري البسيط رقم (٨ ,٣)، كما أن توزيع السكان حسب المناطق الإدارية وتوزيع الجمعيات الخيرية حسب المدن، وتوزيع الأدواء حسب الأقطار وما شابه ذلك من أمثلة تعتبر كلها ضروبا من التصنيف الجغرافي.

وتنطبق جميع أنواع التصنيف السالفة الذكر على الجداول المركبة كما هي الحال مع الجداول البسيطة. ويمكن للقارئ أن يجترح لنفسه وبنفسه أمثلة لجميع التصانيف أعلاه لجداول مزدوجة كما يمكنه النظر إلى الجدول المزدوج رقم (٩ ,٣) والذي يقف مثالا للتصنيف النوعي، وإلى الجدول المزدوج رقم (١٢ ,٣) الذي يمثل التصنيف الكمي.

٣,٣,٣ الجداول الإحصائية البسيطة

تمثل الجداول ذات الأرقام من (٣ ,٣) إلى (٨ ,٣) أمثلة للجداول الإحصائية البسيطة.

الجدول رقم (٣,٣). توزيع خريجي القرية (أ) حسب نوع الشهادة.

نوع الشهادة	العدد (التكرار)
دكتوراه	٣
ماجستير	١٢
بكالوريوس	٣٠
المجموع	٣٥

الجدول رقم (٣,٤). الحجاج القادمون إلى السعودية في سنوات مختارة.

السنة	(عدد الحجاج (التكرار)
١٣١٥هـ	٤٩٥١٧
١٣٦٠هـ	٢٣٨٦٣
١٣٦٥هـ	٦١٢٨٦
١٣٧٠هـ	١٠٠٠٥٧٨

المصدر: الكتاب الإحصائي السنوي للمملكة العربية السعودية (١٣٨٦هـ = ١٩٦٦م)

الجدول رقم (٣,٥). توزيع ٤٠ مصلّيا حسب عدد الصلوات التي يؤدونها يوميا مع الجماعة.

عدد الصلوات في جماعة/اليوم	عدد المصلين (التكرار)
صفر	٣
١	٥
٢	٩
٣	١٢
٤	٨
٥	٣
المجموع	٤٠

الجدول رقم (٣,٦). توزيع مئة رجل أعمال ممن يبلغون ٤٠ سنة من العمر حسب عدد مرات السفر إلى الخارج.

عدد مرات السفر	أقل من ٥	٥-٩	١٠-١٤	١٥-١٩	٢٠-٢٤	٢٥ فأكثر	المجموع
عدد رجال الأعمال المسافرين (التكرار)	٨	١٢	١٨	٣١	٢٥	٦	١٠٠

الجدول رقم (٣,٧). توزيع خمسين طالبا تبعا لفئة العلامات التي تحصلوا عليها في اختبار لمادة الإحصاء الاجتماعي.

فئة الدرجات	٢٠-٢٩	٣٠-٣٩	٤٠-٤٩	٥٠-٥٩	٦٠-٦٩	٧٠-٧٩	٨٠-٨٩	المجموع
عدد الطلاب (التكرار)	٤	٦	١٢	١٤	٩	٣	٢	٥٠

الجدول رقم (٣,٨). الدول التي أنتجت أكثر من مئة مليون طن متري من البترول الخام في عام ١٩٦٩م

الدولة	الإنتاج (مليون طن متري)
الولايات المتحدة	٥١٠
روسيا	٣٢٨,٨
فنزويلا	١٨٧
إيران	١٦٧
ليبيا	١٥٠
السعودية	١٤٨
الكويت	١٣٩

المصدر : عالم النفط ، بيروت ، ١٠ يناير ١٩٧٠م

٣,٣,٤ الجداول الإحصائية المركبة

تمثل الجداول ذات الأرقام من (٣ , ٩) إلى (٣ , ١٢) أمثلة للجداول الإحصائية المركبة .

الجدول رقم (٣,٩). توزيع عشرين شخصا حسب متغيري الجنس والحالة التعليمية.

الحالة التعليمية / الجنس	متعلم	غير متعلم	المجموع
ذكر	٧	٣	١٠
أنثى	٥	٥	١٠
المجموع	١٢	٨	٢٠

الجدول رقم (٣,١٠). توزيع مائة رجل تبعا لمتغيري الحالة الزواجية وأسلوب المعيشة.

الحالة الزواجية / أسلوب المعيشة	متزوج	عزب	مطلق	أرمل	المجموع
حضري	١٥	٢٥	١٠	٥	٥٥
ريفي	٢٠	١٥	٣	٧	٤٥
المجموع	٣٥	٤٠	١٣	١٢	١٠٠

الجدول رقم (٣,١١). حجم النجاح بالثانوية في المدينة (أ) حسب حالة النجاح والجنس ونوع التعليم.

حالة النجاح	ناجحون بدون رسوب		ناجحون مع رسوب في مادة		ناجحون مع رسوب في مادتين		
الجنس / نوع التعليم	بنين	بنات	بنين	بنات	بنين	بنات	المجموع
حكومي	٦٧٠٠	٣٠٠٠	١٧٠٠	٦٧٠	٦٠٠	٢٣٠	١٢٩٠٠
أهلي	١٣٠٠	٦٠٠	٣٤٠	١٣٤	١٢٠	٤٦	٢٥٤٠
المجموع	٨٠٠٠	٣٦٠٠	٢٠٤٠	٨٠٤	٧٢٠	٢٧٦	١٥٤٤٠

الجدول رقم (٣,١٢). توزيع عدد من النساء اللائي سبق لهن الزواج حسب متغير العمر والعمر عند الزواج الأول.

فئة العمر / فئة العمر عند الزواج الأول	١٠-١٤	١٥-١٩	٢٠-٢٤	٢٥-٢٩	٣٠-٣٤	٣٥-٣٩	٤٠-٤٥	المجموع
١٠-١٤	٧	-	-	-	-	-	-	٧
١٥-١٩	١٠	٢٥	-	-	-	-	-	٣٥
٢٠-٢٤	١٥	٢٠	١٩	-	-	-	-	٥٤
٢٥-٢٩	٢٤	٣٠	١٥	١٠	-	-	-	٧٩
٣٠-٣٤	٢٩	٣٤	٢٥	١٥	١٠	-	-	١١٣
٣٥-٣٩	٣٠	١٨	١٦	١٥	٩	٣	-	٩١
٤٠-٤٥	٤٠	٢٠	١٠	٣	-	-	-	٧٣
المجموع	١٥٥	١٤٧	٨٥	٤٣	١٩	٣	-	٤٥٢

٣,٣,٥ عملية جدولة البيانات Tabulation Process

بعد أن تأملنا مظهر جداول الإحصاءات التي تتم جدولتها ممهدين إلى ذلك باستعراض أنواع الجداول الإحصائية وأسس تصنيف المعلومات عند جدولتها، ندلف

الآن إلى تناول عملية الجدولة نفسها وكيف تتم بعد حصولنا على القراءات الخام. ولهذا الغرض سوف نقوم بتقسيم البيانات أولا إلى بيانات وصفية وبيانات كمية ثم نقوم ثانيا بتصنيف البيانات الكمية إلى بيانات كمية ذات مدى[١] صغير نسبيا وأخرى ذات مدى كبير نسبيا.

١,٥,٣,٣ جدولة البيانات الوصفية [ظاهرة واحدة]

لوصف كيفية جدولة البيانات الاسمية دعنا نستعيد المثال (٢ , ٣) الذي تطرقنا إليه في بداية هذا الفصل والذي يتلخص في أن البيانات التالية تمثل عينة من ٢٠ (عشرين) موظفا بإحدى الإدارات الحكومية عرضوا تبعا لمسمى الحالة الزواجية لكل منهم كما يلي:

متزوج، عزب، متزوج، متزوج، عزب
مطلق، أرمل، عزب، مطلق، أرمل
متزوج، متزوج، عزب، أرمل، متزوج
مطلق، متزوج، مطلق، عزب، أرمل

فإذا طلب منا القيام بإنشاء توزيع تكراري بسيط[٢]، ما علينا إلا اتباع الخطوات التالية: نبدأ برسم جدول يتكون من ثلاثة أعمدة (أو خانات) يخصص أحدها (الأول من جهة اليمين) لمتغير الحالة الزواجية ويخصص آخر (الثاني من جهة اليمين) لعلامات الحصر حيث توضع علامات تصنيف (في شكل خطوط مائلة) كل موظف أمام فئة الحالة الزواجية التي تناسبه من بين الفئات الأربع المرصودة في البيانات الخام وهي متزوج، عزب، مطلق، أرمل والتي يتم تدوينها في العمود الأول كما يبدو للناظر في الجدول رقم (١٣ , ٣). ويتم وضع علامات تصنيف كل موظف تبعا لفئة حالته الزواجية بأن نأخذ الحالة الزواجية لكل موظف الواحد تلو الآخر (أفقيا أو عموديا وفق أحد الترتيبين للبيانات المعطاة) ونضع شرطة مائلة هكذا (/) لكل حالة نأخذها أمام الصفة (أي، الفئة) المناظرة لها وذلك في العمود الثاني، أي عمود علامات التصنيف. وإذا

(١) المدى هو الفرق بين أكبر قراءة وأصغر قراءة من بين القراءات المعطاة.
(٢) كثيرا ما يطلق على الجدول التكراري اسم «التوزيع التكراري».

زادت مثل هذه الخطوط المائلة على أربعة خطوط أمام فئة معينة نضع الشرطة (الخط المائل) الخامسة على صورة خط مائل عكسي يقطع الخطوط الأربعة السابقة فنحصل على ما يسمى بـ«الحزمة» [انظر الجدول رقم (١٣ , ٣)] قبل أن نسترسل في وضع الشرطات الأخرى بنفس الطريقة الآحادية إذا زادت الشرطات عن الخمس أمام الفئة الواحدة. . وهكذا إلى أن نستوعب جميع الموظفين تبعا للحالة الزواجية لكل واحد منهم بهذه الكيفية. بعد ذلك نقوم بترجمة هذه العلامات (أو الخطوط المائلة أو الشرطات) إلى أرقام بحيث نضع أمام كل فئة (أي، حالة زواجية) عدد التكرارات وذلك بصورة رقمية حيث تحتل هذه الأرقام العمود الثالث والأخير كما هو واضح في الجدول رقم (١٣ , ٣) والذي يطلق عليه «جدول تفريغ البيانات».

الجدول رقم (٣,١٣). تفريغ بيانات المثال (٣,٢)

الحالة الزواجية	العلامات	التكرارات
متزوج	~~////~~ //	٧
عزب	~~////~~	٥
مطلق	////	٤
أرمل	////	٤
المجموع	-	٢٠

ونخلص من جدول تفريغ البيانات (١٣ , ٣) إلى الجدول التكراري البسيط رقم (١٤ , ٣) والذي يتكون فقط من خانتين اثنتين تلخيصا للظاهرة محل الدراسة وهي الحالة الزواجية في هذا المثال. ويلاحظ القارئ أن الفرق الوحيد بين الجدولين رقمي (١٣ , ٣) و(١٤ , ٣) هو أن الأول يحتوي على عمود علامات التفريغ أو علامات الحصر بينما يتخلص الثاني من هذا العمود.

الجدول رقم (٣,١٤). التوزيع التكراري لعشرين موظفا وفق متغير الحالة الزواجية [جدولة بيانات المثال «٣,٢»].

فئات	التكرار
متزوج	٧
عزب	٥
مطلق	٤
أرمل	٤
المجموع	٢٠

٢,٥,٣,٣ جدولة البيانات الوصفية [ظاهرتان]

كما قد تعلمنا عند تناول الحديث عن أنواع الجداول التكرارية فإن الجدول التكراري المزدوج أو المركب يتم إنشاؤه من بيانات خام تتضمن صفتين (أو أكثر) لكل حالة دراسية فيما لو كانت البيانات الخام بيانات اسمية. وبمعنى آخر، فإن هناك لكل حالة مدروسة «اقتران» يتم تعيينه بصفتين متلازمتين يحكمهما وضع البيانات الخام كما هو الحال في المثال (٣,٣) التالي الذي يعرض ٢٠ (عشرين) شخصا تبعا لمتغيري الحالة التعليمية والجنس لكل واحد منهم.

مثال (٣,٣)

متعلم، ذكر	متعلم، ذكر	أمي، أنثى	متعلم، ذكر
أمي، أنثى	أمي، ذكر	متعلم، أنثى	أمي، أنثى
متعلم، ذكر	متعلم، ذكر	أمي، ذكر	أمي، ذكر
متعلم، ذكر	متعلم، أنثى	متعلم، أنثى	متعلم، أنثى
أمي، أنثى	متعلم، ذكر	أمي، ذكر	متعلم، ذكر

ولتفريغ بيانات المثال (٣ , ٣) في جدول تفريغ ثم الخلوص من هذا إلى جدول تكراري مزدوج نتبع نفس الخطوات التي تم اتباعها في البند «١ , ٥ , ٣ , ٣» الخاص بجدولة البيانات الوصفية التي تتضمن متغيرا واحدا (أي، ظاهرة واحدة) مع الأخذ في الاعتبار، هذه المرة، أن وضع الشرطات على سمات البيانات الخام ينتهج سبيلا يضع الخط على زوج من الصفات بدلا من صفة واحدة (انظر المثال). وعلى هذا النهج سوف نتوصل إلى جدول التفريغ رقم (١٥ , ٣) ومنه نخلص إلى الجدول التكراري المزدوج رقم (١٦ , ٣) والذي يعكس جدولة البيانات الخام في المثال (٣ , ٣).

الجدول رقم (٣.١٥). تفريغ بيانات المثال (٣.٣).

الحالة التعليمية	متعلم		أمي		
الجنس	العلامات	التكرار	العلامات	التكرار	المجموع
ذكر	~~||||~~ |||	٨	||||	٤	١٢
أنثى	||||	٤	||||	٤	٨
المجموع	-	١٢	-	٨	٢٠

الجدول رقم (٣.١٦). التوزيع التكراري لعشرين شخصا حسب متغيري الحالة التعليمية والجنس [جدولة بيانات المثال (٣.٣)].

الحالة التعليمية / الجنس	متعلم	أمي	المجموع
ذكر	٨	٤	١٢
أنثى	٤	٤	٨
المجموع	١٢	٨	٢٠

٣,٥,٣,٣ جدولة البيانات الكمية «المنفصلة» Discrete ذات المدى الصغير نسبيا[٣]

لقد تعلمنا أن البيانات الكمية تأخذ أرقاما عددية وليس صفات كما ينطبق ذلك على البيانات النوعية أو الوصفية. وعرفنا أيضا أن البيانات الكمية تنقسم بدورها إلى بيانات كمية منفصلة discrete وبيانات كمية متصلة continuous بنفس القدر الذي عرفنا فيه أن الفرق بين النوعين من البيانات هو: أن مقادير النوع الأول (أي المنفصلة) لا تحتمل التجزئة، مثل عدد الناس، بينما تحتمل ذلك مقادير النوع الثاني (المتصلة)، مثل الدخل الشهري للفرد. ولأن فلسفة الجدولة لا تختلف في الأسس العامة من نوع إلى آخر من نوعي البيانات الكمية وإنما تختلف اختلافا طفيفا في بعض التفاصيل المتعلقة بتحديد الفئات وأطوالها كما سوف يتضح باتصال الطرح، فسوف نكتفي بتناول جدولة البيانات الكمية «المنفصلة» ذات المدى الصغير نسبيا بالإضافة إلى جدولة البيانات الكمية «المتصلة» ذات المدى الكبير نسبيا تاركين عملية جدولة البيانات «المنفصلة» ذات المدى الكبير نسبيا وتلك التي تتعلق بالبيانات «المتصلة» ذات المدى الصغير نسبيا لأمر التمارين في نهاية هذا الفصل.

المثال (٣,٤) يطرح عرضا لبيانات خام تتعلق بتدوين (٢٠) أسرة تبعا لعدد أفراد كل منها، أي تبعا لحجم كل أسرة من هذه الأسر العشرين.

مثال (٣,٤). أحجام عشرين أسرة.

٢	٢	٣	٥	٦	٤	٤
٣	٥	٦	٤	٤	٥	٢
٣	٤	٥	٦	٤	٥	

(٣) ليست هنالك قاعدة ثابتة يتم الرجوع إليها لتحديد الفرق بين ما يمكن اعتباره مدى صغيرا أو مدى كبيرا، وإنما يعتمد ذلك بالدرجة الأولى على الحس الإحصائي statistical sense لمن يتولى جدولة البيانات الخام. ولكن يمكن عموما أن يسترشد بإمكانية اعتبار المدى قصيرا إذا كان الفرق بين أعلى وأدنى قراءة لا يتعدى بضع حالات (أي بين ٥ إلى ١٠ حالات)، واعتباره مدى كبيرا إذا تعدى الفرق هذه الحدود.

ولتكوين جدول تكراري لبيانات المثال (٤ , ٣) نقوم أولا بحساب المدى لمعرفة ما إذا كان يجدر اعتباره مدى صغيرا نسبيا أو كبيرا نسبيا. وبحساب المدى باستخدام القاعدة: المدى = أكبر قراءة - أصغر قراءة، يتضح أن قيمته تساوي ٤ (٦-٢=٤)، وهي قيمة صغيرة نسبيا مما يستدعي وضع تسلسل الأسر من الأصغر إلى الأكبر أو العكس في عمود يشغله متغير «حجم الأسرة» دونما لجوء إلى تحديد فئات ببداية «دنيا» ونهاية «عليا» كما سوف يتضح أنه الحال مع عملية تفريغ البيانات الكمية ذات المدى الكبير نسبيا عندما نتقدم في الطرح. وكما هو مألوف، فإن جدول التفريغ والجدول التكراري يبدوان كما في الجدولين رقمي (١٧ , ٣) و(١٨ , ٣). ويلاحظ أننا نتحصل على عدد فئات الأحجام التي يتم توزيع جميع الأسر عليها باستخدام المعادلة: عدد الفئات = المدى الكلي + ١، أي: (٦-٢) + ١ = ٥ وهو عدد فئات توزيعنا التكراري في الجدول رقم (١٨ , ٣) تسلسلا من أصغر قيمة من بين القيم الخام (البيانات غير المنظمة) إلى أكبر قيمة فيها. وسوف نرى أن هذا النهج في تحديد عدد الفئات يختلف عن نظيره المتبع في حالة البيانات الكمية ذات المدى الكبير.

الجدول رقم (٣.١٧). تفريغ بيانات المثال (٣.٤).

حجم الأسرة	العلامات	التكرارات
٢	///	٣
٣	///	٣
٤	~~////~~ /	٦
٥	~~////~~	٥
٦	///	٣
المجموع	-	٢٠

الجدول رقم (٣.١٨). التوزيع التكراري لعشرين أسرة حسب متغير حجم الأسرة [جدولة بيانات المثال (٣.٤)].

حجم الأسرة	التكرار
٢	٣
٣	٣
٤	٦
٥	٥
٦	٣
المجموع	٢٠

٣.٣.٥.٤ جدولة البيانات الكمية «المتصلة» ذات المدى الكبير نسبيا

أ – تمهيد

إن إنشاء التوزيع التكراري للأعداد الكبيرة من البيانات الخام يتطلب انتباها أكبر مما هو مكرس تجاه نظيره المتعلق بالأعداد الصغيرة ذات المدى الصغير. ذلك أن تحديد عدد الفئات المناسب وتحديد أطوال هذه الفئات(٤) يحتاج إلى حس إحصائي معقول حتى لا ننتهي إلى شبه تلخيص لهذه البيانات لا يساعدنا على استخلاص صورة واضحة عن المجتمع الذي هو قيد البحث بأخذنا لعدد كبير من الفئات بأطوال قصيرة، أو إلى تلخيص مُبْتَسر لهذه البيانات بأخذنا لعدد صغير منها بأطوال مديدة فتضيع منا بذلك ملامح الظاهرة المدروسة في اختصار مخل وغير فصيح. لذلك فإننا سوف نتبع خطوات محددة تقودنا إلى أمثل السبل لتكوين توزيع تكراري للبيانات الخام ذات المدى الكبير. ولكي يكون تصور هذه العملية أيسر للذهن، دعنا نأخذ مثالا تطبيقيا يساعدنا على ذلك. المثال (٣, ٥) يعرض عدد الأيام التي قضاها خمسون حدثا في إحدى الدور

---

(٤) طول الفئة هو الفرق بين بداية الفئة ونهايتها بمعالجة خاصة. فمثلا إذا نظرت إلى الجدول رقم (٣, ٧) ص(٤٩) يكون طول الفئة الأولى ٢٠–٢٩ هو (٢٩+١) – ٢٠، أي ١٠.

الإصلاحية للأحداث المنحرفين. لاحظ أن تدوين الأيام تم في صورة مربع طول ضلعه سبعة نقاط (أي سبعة تدوينات) إلا أن أحد أضلاعه (أقصى يمين المدونة) يحوي نقطة ثامنة في أسفله (النقطة ٤٨). واتبع هذا النهج فقط لتسهيل عملية الحصر أفقيا أو رأسيا وليس لأي سبب آخر يتعلق بالكيفية التي بها تدون البيانات الخام.

مثال (٣,٥). عدد الأيام التي قضاها خمسون حدثا بإحدى الدور الإصلاحية.

~~٢٥~~	٣٧	~~٢٩~~	~~٢٨~~	~~٢٥~~	~~٢٦~~	~~٢٦~~
٣٤	٣٠	٣٠	٣٣	٣٤	٣٢	٣١
٣١	٣١	٣٢	٣١	٣٣	٣٠	٣٤
٣٠	٣٢	٣٩	٣٥	٣١	٣٣	٣١
٣٥	٣٧	٣٨	٣٧	٣٥	٣٩	٣٦
٣٨	٤٠	٣٨	٤٣	٣٧	٤٤	٣٦
٤٣	٣٩	٤٤	٤٠	٤٠	٤٥	٤٩
٤٨						

وبالإضافة إلى البيانات المثبتة في المثال (٣,٥) يمكن إيراد كثير من أمثلة البيانات الكمية المتصلة ذات المدى الكبير كتلك التي تعكس الدخول الشهرية لعدد كبير من الأسر، أو علامات الطلاب في اختبار لمادة ما، أو أوزان أو أطوال عدد كبير من الأطفال الذين هم تحت الدراسة، إلخ. ويتطلب أمر تبويب البيانات الكمية ذات المدى الطويل تقسيمها إلى أقسام تسمى «فئات» أو «فترات» intervals. ويبين التوزيع التكراري كيفية توزيع المشاهدات أو الحالات المدروسة على كل فئة من الفئات. ولإنشاء التوزيع التكراري يتعين تحديد المدى الكلي للقيم المعطاة، حيث يتم تقسيم هذا المدى إلى عدد مناسب من الفئات (يتراوح عادة بين ٥ إلى ١٥ فئة) بأطوال منتظمة (أي متساوية من فئة إلى أخرى) في الغالب الأعم. ويكون لكل فئة من هذه الفئات حد أدنى وآخر أعلى (أو بداية ونهاية) يعملان على تحديد طولها أيضا كما سوف يتضح من التقدم في سياق الطرح. كما أن لكل فئة من الفئات مركز معين يتم الحصول عليه بإضافة الحدين

الأدنى والأعلى وتقسيم الناتج على ٢ (اثنين). ولمركز الفئة أهمية بالغة في التحليل الإحصائي كما سوف يتضح في الفصول اللاحقة.

ب - خطوات تكوين التوزيع التكراري

لإنشاء جدول تكراري لبيانات كمية ذات مدى كبير هناك خطوات رئيسة يتم اتباعها تسلسلا كما يلي:

١- تحديد عدد الفئات المناسب.
٢- تحديد طول الفئة المناسب.
٣- إنشاء الفئات.
٤- حصر عدد القيم التي تناظر كل فئة ووضعها أمام تلك الفئة.

رغم أن هناك عدة وسائل حسابية تم اقتراحها لتحديد أنسب عدد للفئات إلا أنه سيتم الاكتفاء هنا بالتطرق إلى أنجعها وهي قاعدة ستيرجس Sturges' Rule التي يمكن تلخيصها في المعادلة:

$$\text{م} = ١ + ٣,٣٠ \text{ لو}_{١٠} (\text{ن})$$

حيث:

م = العدد المناسب للفئات.
ن = عدد المشاهدات أو الحالات المدروسة.
لو = اللوغريثمات العادية[٥]، حيث «لو» مختصر لـ«لوغريثم» وبتطبيق هذه القاعدة لإيجاد عدد الفئات المناسب من بيانات المثال (٣ , ٥) نعوض قيمة ن في المعادلة فيصبح عدد الفئات المناسب:

$$\text{م} = ١ + ٣,٣٠ \text{ لو}_{١٠} (٥٠)$$
$$= ١ + ٣,٣ \times ١,٦٩٩$$
$$= ١ + ٥,٦٠٧$$

---

(٥) اللوغريثمات العادية هي لوغريثمات الأساس ١٠ (لو$_{١٠}$) ← Log_{10} وهي تختلف عن اللوغريثمات الطبيعية (لو$_{هـ}$) ← Log_e.

= ٦,٦٠٧

= ٧ فئات تقريبا .

ويجب التنبيه إلى أن هذه القاعدة هي مجرد وسيلة من الوسائل للحصول على العدد التقريبي للفئات وليس هناك أدنى إلزام لمبوب البيانات للأخذ بها دائما لأنه قد يرى بحسه الإحصائي أن العدد الأنسب للفئات يكون أقل أو أكثر بقليل بما تقرره نتيجة هذه المعادلة . ولتأكيد هذا الاتجاه سوف نثبت عدد الفئات في مثالنا هذا في العدد «٥» الذي هو أقرب لمضاعفات «الخمسة» و«العشرة» الأريح للذهن بدلا من العدد ٧ . أما بالنسبة للذين ليس لديهم إلمام باللوغريثمات (وإن كانت الاستعانة بالحاسبات الإلكترونية الصغيرة تكفي لإيجاد لوغريثم أي عدد) فيمكنهم الاسترشاد بالجدول رقم (١٩ , ٣) والذي هو نتاج لتطبيق قاعدة ستيرجس (مع التقريب لأقرب عدد صحيح) لمعرفة العدد التقريبي المناسب لفئات مختلف أعداد المشاهدات .

الجدول رقم (٣,١٩). عدد الفئات المناسب المناظر لعدد المشاهدات طبقا لقاعدة ستيرجس.

عدد المشاهدات	٣٠	٥٠	١٠٠	٢٠٠	٥٠٠	١٠٠٠	٢٠٠٠	٥٠٠٠	١٠٠٠٠	٢٠٠٠٠	٣٠٠٠٠	٥٠٠٠٠	١٠٠٠٠٠
عدد الفئات	٦	٧	٨	٩	١٠	١١	١٢	١٣	١٤	١٥	١٦	١٧	١٨

وبحصولنا على عدد الفئات المناسب نتدرج في خطواتنا لإنشاء التوزيع التكراري للمثال (٥ , ٣) فنتناول الخطوة (٢) وهي تحديد طول الفئة المناسب (ط) .

ولتحديد الطول المناسب للفئة نستخدم القاعدة :

$$\text{ط} = \frac{\text{المدى الكلي}}{\text{عدد الفئات (م)}}$$

$$= \frac{\text{القيمة الكبرى} - \text{القيمة الصغرى}}{\text{عدد الفئات (م)}}$$

$$= \frac{٤٩ - ٢٥}{٥}$$

$= \frac{٢٤}{٥} =$ ٤,٨ = ٥ تقريبا

ويلاحظ أننا قمنا بجبر الكسر عند القسمة إلى الواحد الصحيح وهذا هو ما ينبغي فعله مع نتيجة أي عملية قسمة مماثلة اتخذت عددا كسريا مهما كانت قيمة الكسر. ولإنشاء الفئات الخمس (الخطوة التالية) التي ننتقل بعدها إلى الخطوة الرابعة والأخيرة، نختار أصغر قيمة في البيانات الخام المعطاة [وهي القيمة «٢٥» في بيانات المثال (٥, ٣)] لتكون بداية الفئة الأولى، ثم نضيف إليها طول الفئة الذي وجدناه [وهو ٥ في المثال (٥, ٣)] للحصول على بداية الفئة الثانية، ولهذه نضيف أيضا طول الفئة الثابت [وهو «٥» في مثالنا] لإيجاد الحد الأدنى (أي البداية) للفئة الثالثة، وهكذا إلى أن ننشئ الحدود الدنيا لجميع فئاتنا. ولإيجاد الحدود العليا للفئات (أي نهايات الفئات) نطرح العدد «١» من بداية الفئة الثانية للحصول على نهاية الفئة الأولى، ونكرر طرح العدد «١» من بدايات الفئات الباقية للحصول على نهايات الفئات السابقة لها. بعد ذلك ندون الفئات بحدودها الدنيا والعليا (أي البداية والنهاية لكل فئة) في جدول يتكون من ثلاث خانات (أو أعمدة) تحتل الفئات المنشأة الخانة الأولى (على اليمين) بينما تحتل العلامات الخانة الثانية، أما الخانة الثالثة فتخصص للتكرارات. وبهذا نكون تحصلنا على هيكل جدول التفريغ الذي يمهد لارتياد الخطوة الرابعة والأخيرة وهي حصر القيم ووضعها أمام الفئات بالطريقة التي عهدناها، فيكتمل بذلك إنشاء التوزيع التكراري المطلوب.

ولتوضيح هذه الخطوات عمليا دعنا نطرق المثال (٥, ٣) لنجد

١- تحديد عدد الفئات. عدد الفئات = ٥ (انظر العرض السابق).

٢- تحديد طول الفئة. طول الفئة = ٥ (انظر العرض السابق).

٣- إنشاء الفئات. أصغر قراءة في بيانات المثال (هـ) هي «٢٥»، إذًا:

- بداية الفئة الأولى = ٢٥
- بداية الفئة الثانية = ٢٥ + ٥ = ٣٠
- بداية الفئة الثالثة = ٣٠ + ٥ = ٣٥
- بداية الفئة الرابعة = ٣٥ + ٥ = ٤٠
- بداية الفئة الخامسة = ٤٠ + ٥ = ٤٥

والآن نتحصل على نهايات الفئات الخمس كالتالي :

- نهاية الفئة الأولى = ٣٠ - ١ = ٢٩
- نهاية الفئة الثانية = ٣٥ - ١ = ٣٤
- نهاية الفئة الثالثة = ٤٠ - ١ = ٣٩
- نهاية الفئة الرابعة = ٤٥ - ١ = ٤٤
- نهاية الفئة الخامسة = ٥٠ - ١ = ٤٩

يبقى بعد ذلك تكوين جدول التفريغ رقم (٣,٢٠) والجدول التكراري رقم (٣,٢١) باتباع الخطوات المعهودة فيبدوان كما هما في العرضين التاليين.

الجدول رقم (٣,٢٠). تفريغ بيانات المثال (٣,٥).

فئات الأيام	علامات التفريغ	التكرارات «ك»
٢٥-٢٩	~~||||~~ |	٦
٣٠-٣٤	~~||||~~ ~~||||~~ ~~||||~~ ||||	١٩
٣٥-٣٩	~~||||~~ ~~||||~~ ~~||||~~	١٥
٤٠-٤٤	~~||||~~ ||	٧
٤٥-٤٩	|||	٣
المجموع	-	٥٠

الجدول رقم (٣.٢١). التوزيع التكراري لخمسين حدثا تبعا لفئة الأيام التي قضوها في دور الأحداث.

فئات الأيام	التكرارات «ك»
٢٥-٢٩	٦
٣٠-٣٤	١٩
٣٥-٣٩	١٥
٤٠-٤٤	٧
٤٥-٤٩	٣
المجموع	٥٠

جـ - الحدود الحقيقية للفئات

عندما تكون القراءات كمية متصلة ويمكن بالتالي تجزئة الوحدة منها إلى أجزاء صغيرة ككسور عشرية منها مثل الطول والوزن. . . إلخ، فإن الحد الحقيقي للفئة يمكن أن يأخذ كسرا من الواحد الصحيح مثل الطول ٣ , ٤ بوصة، وعدد الأيام ٥ , ٧ يوما وهكذا. ونلاحظ أننا في الجدول التكراري رقم (٢١ , ٣) قد دونا قراءات الفئات بقيم تقريبية حيث قربت إلى أقرب يوم صحيح قضاه الحدث في دار الملاحظة مثل ٢٥ ، ٢٩ ، ٣٠ وهكذا. ولكن من الممكن أن يكون الحدث قد أمضى ٥ , ٢٤ يوما في الدار مثلا بدلا من ٢٥ يوما كاملا والتي رأينا تقريبها إلى أقرب يوم صحيح لتصير كذلك. وللحصول على القيم الحقيقية لحدود الفئات نطرح «٥ , ٠» (أي نصف الوحدة) كوحدة دقة accuracy unit من الحد الأدنى للفئة ونضيف «٥ , ٠» والتي تمثل وحدة الدقة نفسها إلى الحد الأعلى للفئة ليصبح بالتالي الجدول رقم (٢١ , ٣) كما هو موضح في الجدول رقم (٢٢ , ٣).

الجدول رقم (٣,٢٢). التوزيع التكراري لخمسين حدثا تبعا للفئة «الحقيقية» للأيام التي قضوها في دار الملاحظة.

فئات الأيام (بحدود حقيقية)	التكرار «ك»
٢٤,٥ - ٢٩,٥	٦
٢٩,٥ - ٣٤,٥	١٩
٣٤,٥ - ٣٩,٥	١٥
٣٩,٥ - ٤٤,٥	٧
٤٤,٥ - ٤٩,٥	٣
المجموع	٥٠

د - مراكز الفئات

ذكرنا سابقا أن مركز الفئة يمثل منتصف المسافة بين الحد الأدنى والحد الأعلى للفئة المعينة. ويمكن إيجاد مركز الفئة باستخدام المعادلة:

مركز الفئة = $\frac{١}{٢}$ (بداية الفئة + نهاية الفئة)

ولا تتأثر قيمة مركز الفئة سواء كانت حدود الفئة حقيقية أو تقريبية. وتأتي أهمية مركز الفئة في أننا نفترض دائما أن جميع المشاهدات التي تقع في فئة ما وكأن قيمة كل منها تساوي مركز الفئة تماما. فمثلا، من الجدول رقم (٣,٢٢) نفترض أن كل حدث من الأحداث الستة الذين يقابلون الفئة الأولى (٢٤,٥-٢٩,٥) قضى بدار الملاحظة في المتوسط (سبعة وعشرون) يوما، أي: $\frac{١}{٢}$ (٢٤,٥ + ٢٩,٥) = ٢٧. وهذا الافتراض عبر جميع الفئات يساعدنا كثيرا في إجراء التحليلات الإحصائية، كما سبق وأن ذكرنا، من الجدول التكراري مباشرة دون الرجوع إلى البيانات الخام. وسوف ندرك هذا لاحقا عند نقاش مقاييس النزعة المركزية ومقاييس التشتت.

ولإيجاد مراكز الفئات للجدول رقم (٢٢ , ٣) نقوم هنا في الجدول رقم (٢٣ , ٣) بتثبيت نوعي الفئات الحقيقية منها والتقريبية لتأكيد استخدام أي نوع منهما في إيجاد مراكز الفئات وذلك بمساعدة القاعدة $\frac{١}{٢}$ (بداية الفئة + نهاية الفئة).

الجدول رقم (٣,٢٣). مثال تطبيقي لإيجاد مراكز الفئات.

الحدود التقريبية للفئات	الحدود الحقيقية للفئات	مراكز الفئات «س»	التكرارات «ك»
٢٥-٢٩	٢٤,٥-٢٩,٥	٢٧	٦
٣٠-٣٤	٢٩,٥-٣٤,٥	٣٢	١٩
٣٥-٣٩	٣٤,٥-٣٩,٥	٣٧	١٥
٤٠-٤٤	٣٩,٥-٤٤,٥	٤٢	٧
٤٥-٤٩	٤٤,٥-٤٩,٥	٤٧	٣
المجموع	-	-	٥٠

ويلاحظ من الجدول رقم (٢٣ , ٣) أن الفرق بين مركز فئة ما والذي يليه يساوي ٥ ، وهو طول الفئة لهذا الجدول التكراري المنتظم[٦].

هـ - الجدولان: التكراري النسبي والتكراري النسبي المئوي

التكرارات المطلقة لا تفي بغرض مقارنة الحالات ودرجة تمثيلها للظاهرة التي هي محل الدراسة بطريقة مثلى خاصة عندما يكون عدد التكرارات المطلقة كبيرا إلى حد ما. ولذلك يلجأ الباحثون الاجتماعيون إلى استخدام النسب التي تنسب تكرا. فئة ما إلى المجموع الكلي للتكرارات حتى يتمكنوا من إبراز الأهمية النسبية لمجم

(٦) يسمى الجدول التكراري (أو التوزيع التكراري) جدولا تكراريا منتظما إذا كانت أطوال جميعها متساوية، أي أن طول الفئة ثابت لا يتغير من فئة إلى فئة أخرى.

معينة من الحالات المدروسة مقارنة بمجموع الحالات المدروسة، لكونها - أي الأهمية النسبية - مؤشرا لدرجة شراكة تلك المجموعة في تفسير الظاهرة المدروسة. ولنضرب مثلا لما ذهبنا إليه من واقع إحصاءات الجدول (٢٣ , ٣)، إذ يمكننا القول بأن عدد ٦ (ستة) أحداث (تكرار الفئة الأولى) أمضوا في المتوسط ٢٧ يوما (مركز الفئة الأولى) في دار الملاحظة، من بين مجموعة من (خمسين) حدثا تمت دراسة مدد بقائهم بهذه الدار. ما هي أهمية هؤلاء الستة بالنسبة للخمسين لكي نقيّم بالتالي ثقل وزن المدة المحددة بـ ٢٧ يوما بالنسبة لأطوال مدد البقاء الأخرى؟ سؤال مهم جدا ولكن الإجابة عليه تظل سطحية تماما إن لم نقم بنسبة هؤلاء الستة الأحداث إلى الخمسين لنحدد وزنهم بين المجموعة المدروسة كلها. ولذلك نعتبر إحصائيا أن الخمسين حدثا هم وحدة واحدة ونقوم بنسبة هؤلاء الستة إليهم لنجد التكرر النسبي[٧] لهم كما يلي:

٦ ÷ ٥٠ = ٠,١٢

وبذلك نستطيع القول أنه إذا اعتبرنا أن هؤلاء الخمسين حدثا يمثلون كلا واحدا (أي، يمثلون الوحدة) فإن ١٢ (اثني عشرة) جزءاً من هذا الكل الواحد أمضوا ٢٧ يوما في دار الملاحظة. وكذا يمكن إيجاد الأجزاء الأخرى المقابلة لبقية الفئات، والتي إذا قمنا بجمعها جميعا نجد أنها تساوي «١» صحيح (أي الوحدة). كما يتضح من الجدول (٢٤ , ٣) والذي نطلق عليه اسم الجدول التكراري النسبي.

أما التكرار النسبي المئوي فلا يختلف كثيرا في المفهوم عن التكرار النسبي، غير أننا في الحالة الأولى (أي التكرار النسبي المئوي) نأخذ أي مجموعة من الحالات المدروسة وكأنها تمثل ١٠٠ (مئة) حالة مدروسة، حيث نقوم بعد ذلك بضرب كل تكرار نسبي في ١٠٠ (مئة) لنتحصل على التكرار النسبي المئوي والذي يمثل الأهمية النسبية لكل تكرار مطلق فيما لو كان المجموع الكلي للتكرارات يساوي ١٠٠ (مئة) بدلا من اعتبار المجموع الكلي وحدة واحدة كما في حالة التكرار النسبي. ومن هذا

(٧) التسمية الأصح للتكرار النسبي هي «قيمة التناسب» ولكن يشيع في كثير من كتب الإحصاء تسميتها بـ «النسبة».

الجدول رقم (٣,٢٤). التكرار النسبي للخمسين حدثا.

الفئات	التكرارات	التكرار النسبي
٢٩-٢٥	٦	$\left(\frac{٦}{٥٠}\right) =$ ٠,١٢
٣٤-٣٠	١٩	$\left(\frac{١٩}{٥٠}\right) =$ ٠,٣٨
٣٩-٣٥	١٥	$\left(\frac{١٥}{٥٠}\right) =$ ٠,٣٠
٤٤-٤٠	٧	$\left(\frac{٧}{٥٠}\right) =$ ٠,١٤
٤٩-٤٥	٣	$\left(\frac{٣}{٥٠}\right) =$ ٠,٠٦
المجموع	٥٠	١,٠٠

المنطلق، يمكننا إيجاد التكرار النسبي المئوي للفئة الثالثة، مثلا، بإجراء العملية التالية:

$$\frac{١٥}{٥٠} \times ١٠٠ = ٣٠٪$$

وهذا يعني أنه من بين كل مئة حدث أمضوا عددا من الأيام يتراوح بين ٢٤,٥ إلى ٤٩,٥ في دار الملاحظة نجد أن ٣٠ (ثلاثين) منهم أمضوا في دار الملاحظة عددا من الأيام يتراوح بين ٣٤,٥ إلى ٣٩,٥ [أو، أمضوا ٣٧ يوما في المتوسط (انظر الجدول رقم (٣,٢٣)]. والجدول رقم (٣,٢٥) يوضح التكرار النسبي المئوي لبيانات الخمسين حدثا.

الجدول رقم (٣,٢٥). التكرار النسبي المئوي للخمسين حدثا.

الفئات	التكرارات	التكراري النسبي المئوي (٪)
٢٥-٢٩	٦	$\frac{٦}{٥٠} \times ١٠٠ = ١٢$
٣٠-٣٤	١٩	$\frac{١٩}{٥٠} \times ١٠٠ = ٣٨$
٣٥-٣٩	١٥	$\frac{١٥}{٥٠} \times ١٠٠ = ٣٠$
٤٠-٤٤	٧	$\frac{٧}{٥٠} \times ١٠٠ = ١٤$
٤٥-٤٩	٣	$\frac{٣}{٥٠} \times ١٠٠ = ٦$
المجموع	٥٠	١٠٠

و - الجدولان: التكراري المتجمع الصاعد والتكراري المتجمع الهابط

يكون الاهتمام في كثير من الأحيان منصبا على عدد الحالات أو المشاهدات التي تقابل قيمة ما من قيم الظاهرة التي تمت تجزئتها (أي التي قسمت إلى فئات)، أو تقابل قيمة أقل من قيمة معينة من تلك القيم . ففي مثالنا السابق يمكن أن يطلب منا مثلا إيجاد عدد الأحداث المنحرفين الذين أمضوا أقل من ٣٥ يوما في دار الملاحظة . وللإجابة على هذا السؤال نرتب التكرارات وفق فئات الجدول التكراري البسيط كما يلي مثلا :

الفئات	٢٥-٢٩	٣٠-٣٤	٣٥-٣٩	٤٠-٤٤	٤٥-٤٩
التكرار	٦	١٩	١٥	٧	٣

ومن هذا الجدول يصبح جليا أن عدد الأحداث الذين أمضوا أقل من ٣٥ يوما يمكن الحصول عليه بإضافة تكرار الفئة الأولى إلى تكرار الفئة الثانية، أي: ٦ + ١٩ = ٢٥ حدثا. ويكون هذا التكرار «٢٥» هو التكرار المتجمع الصاعد للفئة الثانية (٣٠-٣٤). أما الطريقة المثلى لإيجاد التكرارات المتجمعة الصاعدة للفئات فتتلخص في أننا نعتمد على الحدود الحقيقية للفئات حيث نكتفي بتدوين الحد الأدنى الحقيقي للفئة ويكتب قبله عبارة «أقل من» في الخانة الأولى للجدول التكراري المتجمع الصاعد ويحتل التكرار المتجمع الصاعد الخانة الثانية لهذا الجدول كما هو واضح في الجدول رقم (٢٦ , ٣)[٨]:

الجدول رقم (٢٦,٣). الجدول التكراري المتجمع الصاعد للخمسين حدثا.

الفئات	التكرار المتجمع الصاعد
أقل من ٢٤,٥	صفر
أقل من ٢٩,٥	٦
أقل من ٣٤,٥	٢٥
أقل من ٣٩,٥	٤٠
أقل من ٤٤,٥	٤٧
أقل من ٤٩,٥	٥٠

وفيما يتعلق بالتكرار المتجمع الهابط فإن الفكرة مشابهة تماما لتلك التي بموجبها كان الاهتمام بالتكرار المتجمع الصاعد، إذ إن الاهتمام ربما كان منصبا على عدد الحالات المدروسة التي تقابل قيما أكثر من قيمة فئوية بعينها. فإذا أردنا مثلا تحديد عدد الأحداث الذين أمضوا أكثر من ٣٩ يوما بالدار، يمكننا أن نستعرض جدولنا التكراري

(٨) لاحظ أن آخر صف في العمود الأول للجدول رقم (٢٦ , ٣) يشير إلى الحد الأدنى الحقيقي للفئة السادسة (التي لم تظهر في جداولنا السابقة) وهذا ما أملته ضرورة حساب التكرار المتجمع الصاعد للفئة الخامسة والأخيرة.

البسيط المعتاد لمثالنا السابق كما هو مثبت أدناه لتسهيل المهمة.

الفئات	٢٥-٢٩	٣٠-٣٤	٣٥-٣٩	٤٠-٤٤	٤٥-٤٩
التكرار	٦	١٩	١٥	٧	٣

ويتضح جليا من هذا الجدول أن عدد الأحداث الذين أمضوا أكثر من ٣٩ يوما بدار الملاحظة هو حاصل جمع تكراري الفئتين الأخيرتين، أي ٧ + ٣ = ١٠. ولإنشاء الجدول التكراري الهابط (أو النازل) نرسم جدولا يتكون من عمودين: الأول يحتوي على سلسلة حدود الفئات حيث يكتفى بالحد الأدنى الحقيقي للفئة ويكتب قبله عبارة «أكثر من» أو «أكبر من»، ويحتوي العمود الثاني على التكرارات المتجمعة الهابطة فيبدو الجدول التكراري المتجمع الهابط كما هو مبين في الجدول رقم (٣,٢٧)[٩]:

الجدول رقم (٣.٢٧). الجدول التكراري المتجمع الهابط للخمسين حدثا.

حدود الفئات	التكرار المتجمع الهابط
أكثر من ٢٤,٥	٥٠
أكثر من ٢٩,٥	٤٤
أكثر من ٣٤,٥	٢٥
أكثر من ٣٩,٥	١٠
أكثر من ٤٤,٥	٣
أكثر من ٤٩,٥	صفر

(٩) لاحظ أن الصف الأخير بالعمود الأول للجدول رقم (٣,٢٧) يتضمن الحد الأدنى الحقيقي للفئة السادسة الذي أملت ضرورة حساب التكرار الهابط للفئة الخامسة تثبيته هنا.

ز – تمثيل الجداول التكرارية للإحصاءات الكمية «المتصلة» بيانيا

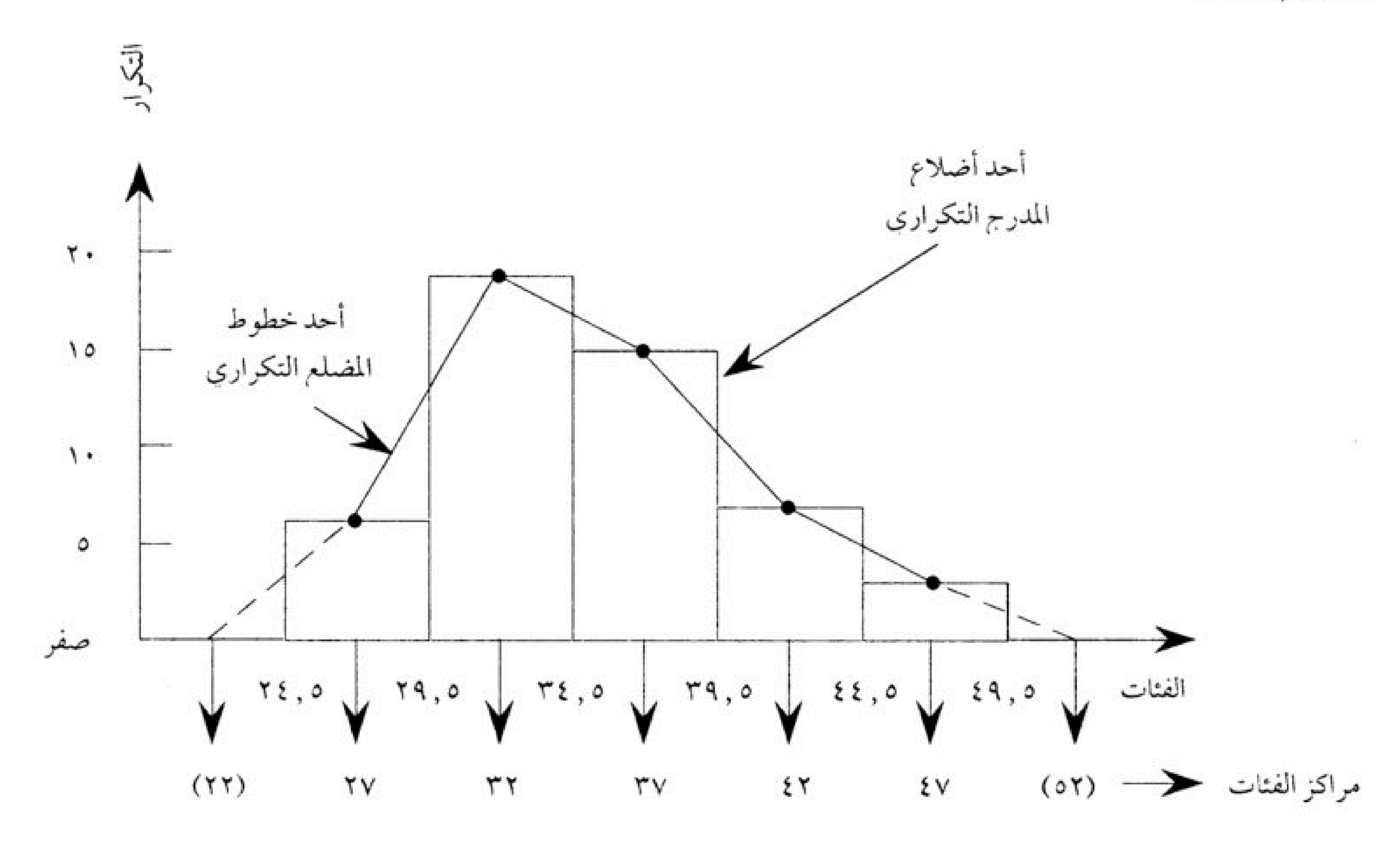
يمكن عرض التوزيعات التكرارية التي تمت مناقشتها سابقا عرضا بيانيا يزيد من إيضاح توزيع الحالات المدروسة على أقسام الظاهرة المختلفة. ويمكننا في جميع أحوال العرض البياني أن نعمل على تمثيل مراكز الفئات والحدود الدنيا والعليا لهذه الفئات في محور أفقي بينما تمثل التكرارات المقابلة لها على المحور الرأسي. ويعتبر المدرج التكراري والمضلع التكراري والمنحنى المتجمع الصاعد والمنحنى المتجمع الهابط من الطرق الشائعة في العرض البياني للبيانات الكمية «المتصلة». وسوف نتناول كلا من هذه الطرق تطبيقا على مثالنا السابق في ما يلي من طرح.

١- المدرج والمضلع التكراريّان: يستخدم كل من المدرج التكراري frequency histogram والمضلع التكراري frequency polygon لعرض بيانات الجدول التكراري البسيط للبيانات الكمية المتصلة وهما مشابهان، على التوالي، للأعمدة البيانية bar graphs [انظر الشكل رقم (١ ,٣أ)] والخط البياني [انظر الشكل رقم (٤ ,٣)] في الفكرة والعرض على وجه العموم إلا أنهما يختلفان في التفاصيل لكون البيانات هنا تعد كمية متصلة وليست كمية منفصلة كما هو الحال في الشكلين السابقين المشار إليهما. ولنبين كيفية رسم المدرج التكراري والمضلع التكراري للبيانات الكمية المتصلة دعنا نعيد رسم الجدول رقم (٢١ ,٣) الخاص بالتوزيع التكراري للخمسين حدثا تبعا لعدد الأيام التي أمضوها في دار الملاحظة والذي يبدو كما يلي:

الفئات	٢٥-٢٩	٣٠-٣٤	٣٥-٣٩	٤٠-٤٤	٤٥-٤٩
التكرار	٦	١٩	١٥	٧	٣

والآن، للاضطلاع برسم المدرج التكراري لهذه البيانات نقوم بوضع الحدود الحقيقية (أو الحدود التقريبية) للفئات على محور أفقي بمقياس رسم معقول، كما نقوم

بوضع التكرارات على محور رأسي بمقياس رسم معقول أيضا، حيث يتقاطع المحوران في نقطة البداية (نقطة الأصل وتأخذ البيان «صفر») كما هو واضح في الشكل رقم (٣,٧). ويرسم فوق كل فئة شكل مضلع (قد يبدو أحيانا على شكل مستطيل وأحيانا أخرى على شكل مربع) تمثل قاعدته طول الفئة وارتفاعه تكرار الفئة. والشكل الناتج عن هذه المضلعات المتلاصقة [انظر الشكل رقم (٣,٧)] يبين المدرج التكراري الذي يعرض التوزيع التكراري الذي يعكسه الجدول رقم (٣,٢١). وتلاصق هذه المضلعات يفصح عن أن نوعية البيانات التي جرى تمثيلها بيانيا هي بيانات كمية «متصلة» وليست «منفصلة».

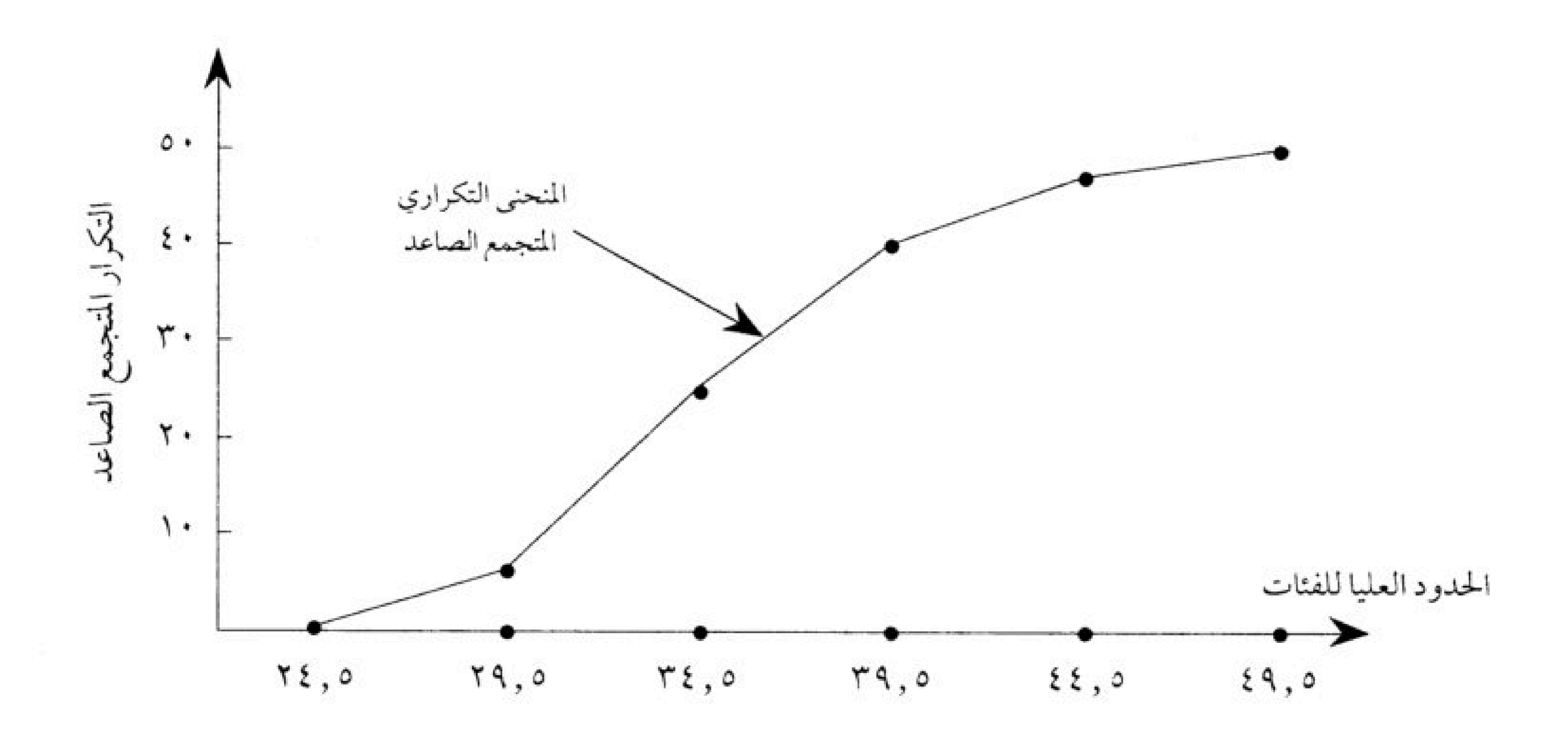


الشكل رقم (٣,٧). المدرج التكراري والمضلع التكراري لبيانات الجدول رقم (٣,٢١).

والشكل رقم (٣,٧) يحمل بين ثناياه أيضا المضلع التكراري الذي تبينه الخطوط المستقيمة المتصلة والتي أنشئت باستخدام مراكز الفئات بدلا من الفئات نفسها كما هو الحال مع المدرج التكراري. ولرسم المضلع التكراري نقوم بتوصيل منتصف قمم المستطيلات (الخاصة بالمدرج التكراري) بخطوط مستقيمة مستخدمين في ذلك المسطرة فيبدو شكل المضلع التكراري كما هو في الشكل رقم (٣,٧). ويلاحظ أن المضلع

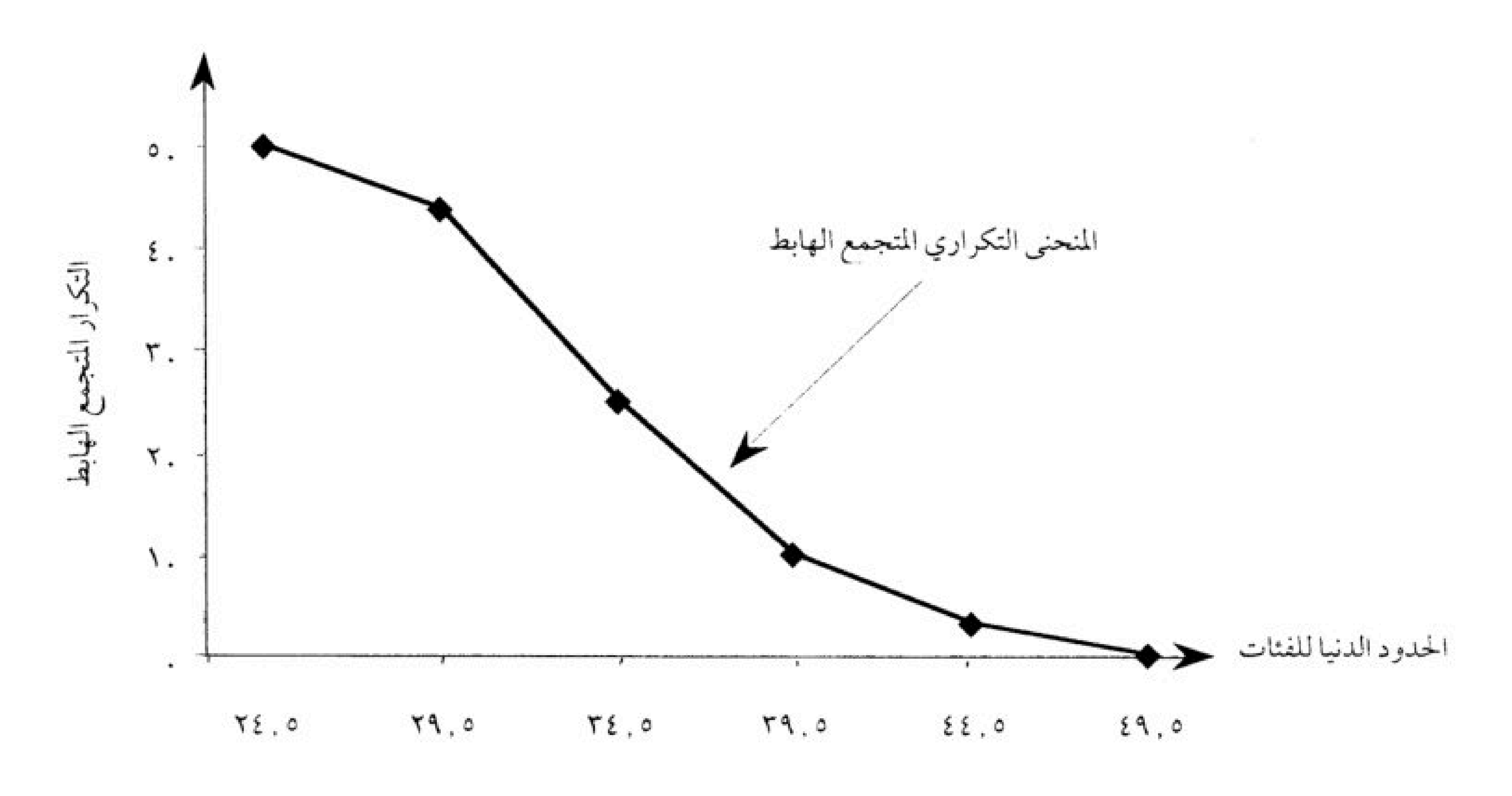
التكراري يبدأ عند مركز الفئة السابقة للفئة الأولى وينتهي عند مركز الفئة التالية للفئة الأخيرة. والخطوط المتقطعة التي تنسب هذين المركزين إلى المضلع التكراري تشير إلى ذلك. ويتم رسم المضلع التكراري بافتراض أن المشاهدات المقابلة لكل فئة تتوزع بانتظام على جميع القيم التي بداخل هذه الفئة. ومما يجدر ذكره أنه لو تم تمهيد الخطوط الواصلة بين مراكز الفئات باليد، يسمى الشكل الناتج عن ذلك «المنحنى التكراري» بدلا من المضلع التكراري.

٢- المنحنيان: المتجمع الصاعد والمتجمع الهابط: لمزيد من الإفصاح عن توزيع الخمسين حدثا وفقا للأيام التي قضوها في دار الملاحظة يمكن تمثيل القراءات التي يحويها كل من الجدول التكراري المتجمع الصاعد [الجدول رقم (٢٦ , ٣)] والجدول التكراري المتجمع الهابط [الجدول رقم (٢٧ , ٣)] تمثيلا بيانيا وذلك برسم المنحنى المتجمع الصاعد والمنحنى المتجمع الهابط لكل منهما على حدة. ولرسم المنحنى المتجمع الصاعد نختط محورين أحدهما عمودي نضع عليه التكرارات المتجمعة الصاعدة بمقياس رسم مناسب، والآخر أفقي نضع عليه نقاط حدود الفئات الحقيقية بمقياس رسم مناسب أيضا. وعند مباشرة وضع النقاط على المحور الأفقي نبدأ بوضع نقطة على هذا المحور عند الحد الأدنى للفئة الأولى (من على اليسار) وكذلك لنبين عدم وجود أية مشاهدات عند هذه النقطة أو قبلها. بعد ذلك نضع نقطة فوق الحد الأعلى للفئة الأولى مباشرة بارتفاع مساو لتكرار هذه الفئة. كذلك نضع نقطة فوق الحد الأعلى للفئة الثانية بارتفاع مساو للتكرار المتجمع الصاعد المناظر لها، وهكذا إلى أن نصل إلى الحد الأعلى للفئة الأخيرة. بعد ذلك يتم توصيل هذه النقاط بخطوط مستقيمة ليكون المنحنى الناتج عن توصيل هذه النقاط ببعضها هو ما نطلق عليه «المنحنى التكراري المتجمع الصاعد» كما هو مبين بالشكل رقم (٨ , ٣) والذي يعكس تمثيل بيانات الجدول التكراري المتجمع الصاعد [الجدول رقم (٢٦ , ٣)] بيانيا باستخدام المنحنى التكراري المتجمع الصاعد.



الشكل رقم (٣,٨). المنحنى التكراري المتجمع الصاعد لبيانات الجدول رقم (٣,٢٦).

ولرسم المنحنى التكراري المتجمع الهابط نتبع طريقة مشابهة لتلك التي اتبعناها عند رسم نظيره المنحنى التكراري المتجمع الصاعد. ولإتمام عملية رسم المنحنى التكراري المتجمع الهابط (أو النازل) لقراءات الجدول رقم (٣ , ٢٧) وبعد أن نختط المحورين المتعامدين بنفس البيانات السابقة، نبدأ بوضع نقطة على المحور الأفقي عند الحد الأعلى للفئة الأخيرة لنبين عدم وجود أية مشاهدات عند هذه النقطة أو بعدها. بعد ذلك نضع نقطة فوق الحد الأدنى للفئة السابقة لها بارتفاع مساو لتكرار هذه الفئة ثم نقطة فوق الحد الأدنى للفئة التي تسبق هذه بتكرار مساو لتكرارها المتجمع النازل، وهكذا حتى نضع نقطة فوق الحد الأدنى للفئة الأولى مباشرة وبارتفاع مساو للمجموع الكلي للتكرارات. عندئذ يتم توصيل هذه النقاط ببعضها بخطوط مستقيمة مثلما فعلنا للحصول على المنحنى التكراري المتجمع الصاعد. ويسمى المنحنى الذي يتم الحصول عليه بتوصيل هذه النقاط «المنحنى التكراري المتجمع الهابط» كما هو مبين بالشكل رقم (٣ , ٩) ترجمة لإحصاءات الجدول رقم (٣ , ٢٧) بيانيا.



الشكل رقم (٣.٩). المنحنى التكراري المتجمع الهابط لبيانات الجدول رقم (٣.٢٧).

٣.٤ أسئلة

١- بيّن وجه الفرق بين الأعمدة البيانية والمدرج التكراري مع التعليل.

٢- بين في نقاط محددة ما الذي نرمي إليه من وراء جدولة البيانات.

٣- فيما يلي عدد السيجارات التي يدخنها يوميا أفراد مجموعة تتكون من ١٤ شخصا:

٥	١٠	٥	١٣	٩	١٠	١٨
٨	٩	١٥	١٤	١٢	١٩	٦

والمطلوب:

(أ) أنشئ التوزيع التكراري للبيانات أعلاه إذا علمت أن طول الفئة يساوي ٥.

(ب) أوجد التكرارين النسبي والنسبي المئوي لتوزيعك.

(جـ) هل يمكن أن يحتمل توزيعك حدودا حقيقية للفئات؟ ولماذا؟

٤- فيما يلي عدد التلاميذ بالصف الأول الابتدائي لـ ٤٠ مدرسة :

٨٠	٩٦	٦٢	٧٣	٩٣	٩٦	٧٥	٣٥	٨٨	٤٢
٧٩	٥٣	٦٢	٨٣	٦٩	٧٣	٥٤	٦٦	٧٦	٥٢
٥٩	٦٧	٨٠	٤٩	٦٥	٥٢	٧٥	٨١	٥٦	٦٩
٧٩	٨٩	٨٢	٩١	٨٧	٧٢	٧١	٤٤	٨٠	٨٨

والمطلوب :

(أ) أنشئ الجدول التكراري للبيانات أعلاه .

(ب) ارسم الأعمدة البيانية لجدولك التكراري .

٥- استخدم بيانات المثال (١ , ٣) المثبت في بداية هذا الفصل للإجابة على الأسئلة التالية :

(أ) كوّن جدولا تكراريا مناسبا لبيانات هذا المثال .

(ب) ارسم المدرج التكراري والمضلع التكراري للتوزيع التكراري الذي تحصلت عليه في «أ» وبين ما هو الفرق بين المضلع التكراري والمنحنى التكراري .

(جـ) أوجد التكرارين المتجمع الصاعد والمتجمع الهابط ثم ارسم منحنييهما التكراريين المتجمعين .

٣,٥ اصطلاحات ينبغي تذكرها

- الأعمدة البسيطة/ الأعمدة المزدوجة/ الأعمدة المجزأة .
- البيانات النوعية/ البيانات الكمية .
- الخط البياني .
- الدائرة البيانية .
- خريطة الشريط .
- تفريغ البيانات .
- الجدول التكراري البسيط .
- الجدول التكراري المزدوج .

- التوزيع التكراري.
- قاعدة ستيرجس.
- المدى الكلي للبيانات الخام.
- الفئة/ حدود الفئة/ مركز الفئة/ الحدود الحقيقية/ الحدود التقريبية.
- التكرار النسبي/ التكرار النسبي المئوي.
- الجدول التكراري المنتظم.
- طول الفئة.
- المدرج التكراري.
- المضلع التكراري/ المنحنى التكراري.
- التكرار المتجمع الصاعد/ التكرار المتجمع الهابط.
- المنحنى التكراري المتجمع الصاعد.
- المنحنى التكراري المتجمع الهابط.

الفصل الرابع

مقاييس النزعة المركزية
Measures of Central Tendency

١,٤ مقدمـــة

لقد تعلمنا في الفصل السابق كيف أن تبويب البيانات (أي تنظيمها بوضعها في جداول تكرارية) وعرضها بيانيا يسهمان وبدرجة كبيرة في إيضاح معالمهما الرئيسة بحيث تغدو تلك البيانات المنظمة أدق وصفا بكثير جدا للظاهرة التي جمعت من أجل دراستها مما لو كانت قد تركت مثلما جمعت خاما بلا إفصاح يذكر عن سماتها العامة. إلا أن أمر التبويب والعرض البياني لا يشكل سوى بداية تمهيدية نحو طرق عملية التحليل الإحصائي للبيانات وتفسيرها وذلك بدءا بحساب قيم خاصة اتفق واضعوا علم الإحصاء على استخدامها كمقاييس لتلخيص السمات البارزة للبيانات التي يجمعها الباحثون. ويضطلع بعض أنواع هذه القيم بتوضيح نزعة بيانات الظاهرة (أو المتغير) الذي هو قيد البحث نحو قيمة مركزية بعينها، فيما يضطلع بعضها الآخر بتوضيح نزعة البيانات المعنية نحو التشتت (أو التباعد أو التباين). ففي الحالة الأولى يطلق الإحصائيون على مثل هذه القيم اسم مقاييس النزعة المركزية - موضوع هذا الفصل - وفي الحالة الثانية يطلقون عليها اسم مقاييس التشتت measures of dispersion. وفي كلا الحالين فإنها قيمةً ما يتم حسابها بطريقة خاصة من البيانات التي تم جمعها لتقوم بتمثيل مجموعة تلك البيانات في درجة تمركزها أو درجة تشتتها.

وسوف نخصص هذا الفصل لدراسة مقاييس النزعة المركزية والتي يطلقون عليها أيضا اسم المتوسطات averages. وتشمل المتوسطات عدة مقاييس أهمها

المتوسط أو الوسط الحسابي arithmetic mean ، والوسيط median ، والمنوال mode . وسوف نحصر دراستنا للمتوسطات في هذه المقاييس الثلاثة فقط لشيوع استخدامها في البحوث الاجتماعية بصورة تكاد تنفي استخدام أي بديل آخر لها من جنسها في مثل هذه الأبحاث . وسوف نتناول في دراستنا لهذه المقاييس كيفية حسابها من البيانات غير المبوبة ungrouped data وكذلك كيفية حسابها من البيانات المبوبة grouped data ، كما سوف نذكر عيوب ومزايا كل واحد منها . أما مقاييس التشتت فسوف نخصص لها الفصل القادم بإذن الله .

٤.٢ الوسط الحسابي (أو المتوسط)[١]

وهو من أكثر مقاييس النزعة المركزية استخداما لفائدته العظمى في وصف البيانات وتلخيصها وإتاحة إمكانية المقارنة بين المجموعات المتماثلة التي يراد إجراء المقارنات بين اتجاهاتها أو نزعاتها نحو قيمة مركزية ، علاوة على سهولة حسابه من البيانات المبوبة وغير المبوبة على حد سواء . والمتوسط مألوف تماما حتى لدى عامة الناس ، إذ كثيرا ما نسمع عن متوسط دخل الفرد في دولة ما ، أو متوسط عمر السكان في قطر ما ، أو متوسط أسعار السلع في موسم ما ، أو متوسط درجة الحرارة في شهر من شهور السنة ، وغير ذلك من مختلف الأمور الحياتية . وفكرة المتوسط بسيطة جدا ، إذ لو كان هناك شخصان حمل أحدهما إلى منزله بعد دوام العمل عشر قطع شوكولاته ليفرقها على أبنائه الخمسة فسوف يقتضي العدل أن يكون نصيب الفرد من هؤلاء الأبناء قطعتين ؛ وإذا حمل الآخر في نفس الوقت إلى أبنائه الخمسة كذلك خمس عشرة قطعة من الشوكولاته فسوف يقتضي العدل في تفريقها عليهم أن يأخذ كل واحد من هؤلاء ثلاث قطع . وهذه مسألة تكاد تكون بديهية ويدركها حتى الشخص العادي . وبمنطق

---

(١) الوسط الحسابي يلائم بصورة خاصة البيانات الممكن قياسها على ميزان القياس الفاصل أو النسبي (راجع الجزء ٥ ,٢ ,٢ بالفصل الثاني) ولا يناسب أبدا البيانات الاسمية أو الرتبية .

الرياضيات، فإننا في كلا الحالين لم نقم بأكثر من قسمة مجموع القيم (مجموعة قطع الشوكولاته) على عدد المشاهدات (عدد الأبناء). فبينما كان نصيب الابن في الحالة الأولى هو $\frac{١٠}{٥}$ = ٢ (قطعتين)، كان نصيبه في الحالة الثانية $\frac{١٥}{٥}$ = ٣ قطع - وهكذا تم حساب الوسط الحسابي في أبسط صوره. وللمقارنة والتفسير يمكننا القول إن درجة رفاهية أبناء الشخص الثاني أعلى في المتوسط من درجة رفاهية أبناء الشخص الأول فيما يتعلق بمدى التلذذ بتناول الشوكولاته في ذلك اليوم.

وبعد هذا المثال البسيط لتوضيح فكرة المتوسط سوف نعمل الآن على التعرف بطريقة منهجية على كيفية حسابه من البيانات غير المبوبة ثم من البيانات المبوبة.

١.٢.٤ حساب المتوسط من البيانات غير المبوبة

بداية يمكن تعريف الوسط الحسابي بأنه «القيمة التي لو أعطيت لكل مشاهدة من المشاهدات المدروسة لكان مجموع مثل هذه القيم مساو لمجموع القيم الأصلية لهذه المشاهدات مهما اختلفت». وسوف يتضح هذا التعريف بصورة أفضل حينما نتناول مثالا تطبيقيا لحساب المتوسط فيما بعد. ويعرف المتوسط رياضيا بأنه يساوي المجموع الكلي للقيم مقسوما على عدد القيم. ومجاراة لسياق هذا التعريف الرياضي، دعنا نرمز للمتغير العشوائي الذي هو قيد البحث بالحرف س، وللقيم التي شاهدناها له هكذا: $س_١$، $س_٢$، $س_٣$، . . . ، $س_ن$؛ حيث إن «ن» هو عدد القيم المشاهدة. يمكن بعدئذ أن نعرف الوسط الحسابي رياضيا كما يلي:

$$\text{الوسط الحسابي} = \frac{س_١ + س_٢ + س_٣ + \ldots + س_ن}{ن} \qquad (١)$$

وهو ما ذهبنا إليه من أن الوسط الحسابي يساوي المجموع الكلي للقيم مقسوما على عدد القيم. ولقد اصطلحت معظم مؤلفات الإحصاء باللغة العربية على الرمز «$\bar{س}$»

(ويقرأ سين شرطة) ليدل على المتوسط، وعلى الرمز «مجـ س» ليدل على المجموع الكلي للقيم. ولهذا يمكن أن تأخذ القاعدة (١) الصورة التالية الأكثر اختصارا:

$$\bar{س} = \frac{مجـ\ س}{ن} \qquad (٢)$$

مثال (٤،١)

لدى سؤالهم عن أعمارهم، أجاب ستة أطفال كما يلي:
٦ ، ٣ ، ٥ ، ٦ ، ١٢ ، ١٠ . أوجدالوسط الحسابي لأعمار هؤلاء الأطفال.

١،١،٢،٤ إيجاد الوسط الحسابي بالطريقة العادية

قبل الحل ندرك من هذه المسألة أن المتغير الذي يدرس هو العمر، وعدد القيم المشاهدة ن هو ٦. وإذا رمزنا للعمر بالحرف «س» فإننا نرمز لقيمه كما يلي:
$س_١ = ٦$ ، $س_٢ = ٣$ ، $س_٣ = ٥$ ، $س_٤ = ٦$ ، $س_٥ = ١٢$ ، $س_٦ = ١٠$.
وللحل نطبق القاعدة (٢):

$$\bar{س} = \frac{مجـ\ س}{ن} = \frac{٦ + ٣ + ٥ + ٦ + ١٢ + ١٠}{٦} = \frac{٤٢}{٦} = ٧ \text{ سنوات}.$$

والآن، لنثبت تعريف المتوسط كما أوردناه في بداية تناول الموضوع بأنه «القيمة التي لو أعطيت . . . التعريف».

$س_١$ = ٧ بدلا من القيمة الأصلية وهي ٦
$س_٢$ = ٧ بدلا من القيمة الأصلية وهي ٣
$س_٣$ = ٧ بدلا من القيمة الأصلية وهي ٥
$س_٤$ = ٧ بدلا من القيمة الأصلية وهي ٦
$س_٥$ = ٧ بدلا من القيمة الأصلية وهي ١٢

$س_٦$ = ٧ بدلا من القيمة الأصلية وهي ١٠

المجموع الكلي = ٤٢ بدلا من مجموع القيم الأصلية = ٤٢

٤,٢,١,٢ إيجاد الوسط الحسابي بطريقة الوسط الفرضي

ويمكن أيضا إيجاد الوسط الحسابي بطريقة مختلفة تعتمد على حساب انحرافات القيم المشاهدة عن أي عدد نختاره ونسميه الوسط الفرضي assumed mean ثم إيجاد المجموع الكلي لهذه الانحرافات بدلا من المجموع الكلي للقيم الأصلية الذي هو أساس الطريقة السابقة. وتُفَضَّل طريقة الوسط الفرضي لإيجاد الوسط الحسابي في حالة البيانات ذات التوزيعات التكرارية الكبيرة بفئات لها حدود دنيا وعليا، حيث تخفف كثيرا من الجهد الحسابي الذي يبذل عند استخدام الطريقة المعتادة. ولكن لا بأس من تطبيق هذه الطريقة على مثالنا السابق للتعرف عليها.

فإذا رمزنا للوسط الفرضي بالحرف «ف» وإلى الانحراف بالحرف «ح»، فبمستطاعنا أن نرمز إلى انحرافات القيم المشاهدة عن الوسط الفرضي بما يلي:

$$ح_١ = س_١ - ف$$

$$ح_٢ = س_٢ - ف$$

$$ح_٣ = س_٣ - ف$$

--- = ------

$$ح_ن = س_ن - ف$$

وهكذا بإمكاننا أن نستخدم القاعدة التالية لإيجاد الوسط الحسابي:

$$\bar{س} = ف + \frac{مجـ ح}{ن} \qquad (٣)$$

وبتطبيق هذه القاعدة على بيانات المثال (١ , ٤) باعتبار أننا قد اخترنا العدد «٥» ليمثل الوسط الفرضي فسوف نتحصل على النتيجة التالية:

$$\bar{\text{س}} = ٥ + \frac{\text{مجـ}\left[(٦ - ٥) + (٣ - ٥) + (٥ - ٥) + (٦ - ٥) + (١٢ - ٥) + (١٠ - ٥)\right]}{٦}$$

$$= ٥ + \frac{١٢}{٦}$$

= ٥ + ٢ = ٧ سنوات .

ونلاحظ أن هذه هي نفس النتيجة السابقة بمعنى أن الوسط الحسابي للأعمار هو ٧ سنوات .

٤،٢،١،٣ حساب الوسط الحسابي المرجح

قد يتطلب الأمر في بعض حالات البحوث الاجتماعية أن يتم دراسة متغير ما أو ظاهرة من الظواهر الاجتماعية ذات أوجه متعددة بإعطاء أهميات متفاوتة لتلك الأوجه قبل حساب قيمة مركزية لأي مشاهدة (أو مجموعة من المشاهدات) تبين مدى أهميتها قياسا بالقيمة المثالية لتلك الظاهرة . ففي مثل هذه الحالات يعتمد على ما يسمى بالوسط الحسابي المرجح weighted arithmetic mean ، حيث تعطى أوزان مختلفة لخصائص المتغير الذي هو قيد الدرس قبل حساب الوسط الحسابي . وسوف نأتي الآن بمثال تطبيقي لذلك .

مثال (٤،٢)

تبعا لنظام تقويم للأعمال المدرسية ، تتألف الدرجة الكاملة لأي مادة دراسية من ١٠٠ (مئة) نقطة فيما تتوزع النقاط بحيث تحتل «المشاركة في الفصل» ٥٪ منها ، «حل الواجب المنزلي» ١٥٪ ، «المواظبة على الحضور» ١٠٪ ، «الاختبارات المباغتة» ١٠٪ ، و«الامتحان النهائي» ٦٠٪ من هذه النقاط . فإذا أريد إيجاد الدرجة النهائية لطالب أحرز في الامتحان النهائي لمادة الإحصاء ٧٠ نقطة فيما أحرز في «المشاركة» و«حل الواجب» و«المواظبة» و«الاختبارات المباغتة» على التوالي ٨٠ ، ٨٢ ، ٧٤ و٦٨ نقطة علما بأن الدرجة الكاملة لأي من هذه الأعمال = ١٠٠ ؛ فإنه يمكن أولا

ترتيب البيانات المعطاة كما في الجدول رقم (١ , ٤).

الجدول رقم (١ , ٤). درجات الطالب وأوزان الأعمال المدرسية.

العمل	الدرجة من ١٠٠ «س»	وزن العمل (٪) «و»
المشاركة في الفصل	٨٠	٥
حل الواجب المنزلي	٨٢	١٥
المواظبة	٧٤	١٠
الاختبارات المباغتة	٦٨	١٠
الامتحان النهائي	٧٠	٦٠
المجموع الكلي	٣٧٤	١٠٠

ولحل هذه المسألة نستخدم قانون إيجاد الوسط الحسابي المرجح والذي مفاده حاصل قسمة مجموع «ضرب الأوزان في القيم» على مجموع الأوزان. وبالرموز مستفيدين من تلك التي ثبتت بالجدول رقم (١ , ٤):

$$\bar{س}_ح = \frac{و_١ س_١ + و_٢ س_٢ + و_٣ س_٣ + \ldots + و_ن س_ن}{و_١ + و_٢ + و_٣ \ldots + و_م}$$

$$\bar{س}_ح = \frac{مج\ و\ س}{مج\ و} \qquad (٤)$$

حيث «و» يرمز إلى وزن الخاصية و«س» يرمز إلى قيمة الخاصية، و $\bar{س}_ح$ يرمز إلى الوسط الحسابي المرجح. ولتسهيل إجراء العمليات الحسابية يجمل بنا أن ننظم عملنا كما في الجدول رقم (٢ , ٤).

الجدول رقم (٤.٢). حساب الوسط الحسابي المرجح لدرجات الطالب [مثال ٤.٢].

العمل	الدرجة من ١٠٠ «س»	وزن العمل (٪) «و»	و س
المشاركة في الفصل	٨٠	٥	٤٠٠
حل الواجب المنزلي	٨٢	١٥	١٢٣٠
المواظبة	٧٤	١٠	٧٤٠
الاختبارات المباغتة	٦٨	١٠	٦٨٠
الامتحان النهائي	٧٠	٦٠	٤٢٠٠
المجموع الكلي	-	١٠٠ = مجـ و	٧٢٥٠ = مجـ و $\bar{س}$

ومن الجدول رقم (٢ , ٤) فإن :

$$\bar{س}_{ح} = \frac{٧٢٥٠}{١٠٠} = ٧٢,٥ \text{ درجة}$$

ويلاحظ أنه إذا تمت مساواة جميع الأعمال في الأهمية فإن درجة الطالب تمثل الوسط

الحسابي العادي الذي يحسب باستخدام القاعدة $\bar{س} = \frac{\text{مجـ س}}{ن}$

$$= \frac{٣٧٤}{٥} = ٧٤,٨ \text{ درجة}$$

٤.٢.٢ حساب المتوسط من البيانات المبوبة

يمكن للبيانات المبوبة أن تكون ذات فئات فردية، أي ليست لها حدود دنيا وأخرى عليا، أو ذات فئات فترية، أي لكل فئة من الفئات حد أدنى وآخر أعلى. وفي كلا الحالين فإن قانون إيجاد الوسط الحسابي ثابت لا يتغير سوى أن القيمة «س» التي يتم ضربها في التكرار «ك» تأخذ مركز الفئة في حالة الفئات الفترية بينما تأخذ قيمة الفئة

الآحادية أو الفردية نفسها في حالة الفئات الفردية . وتقرأ قاعدة إيجاد الوسط الحسابي من البيانات المبوبة كما يلي :

$$\bar{س} = \frac{ك_١ س_١ + ك_٢ س_٢ + ك_٣ س_٣ + \ldots + ك_م س_م}{ك_١ + ك_٢ + ك_٣ \ldots + ك_م}$$

$$\bar{س} = \frac{\text{مجـ ك س}}{\text{مجـ ك}} \qquad (٥)$$

مثال (٣, ٤)

أوجد متوسط حجم الأسرة مستخدما بيانات الجدول رقم (١٨ , ٣) بالفصل الثالث . ولتسهيل الحل نضع هذه البيانات في الصورة الموجودة في الجدول رقم (٣ , ٤) مع إضافة البيانات اللازمة لحساب المتوسط .

الجدول رقم (٣, ٤). حساب المتوسط لفئات فردية.

حجم الأسرة «س»	التكرار «ك»	ك س
٢	٣	٦
٣	٣	٩
٤	٦	٢٤
٥	٥	٢٥
٦	٣	١٨
المجموع	٢٠ = مجـ ك	٨٢ = مجـ ك س

من الجدول رقم (٣ , ٤) وباستخدام القاعدة (٥) فإن :

$$\bar{س} = \frac{\text{مجـ ك س}}{\text{مجـ ك}} = \frac{٨٢}{٢٠} = ١ , ٤ \text{ فردا}$$

أي أن متوسط حجم الأسرة لهذه المجموعة من الأسر = ١ , ٤ شخصا .

مثال (٤,٤)

أوجد متوسط عدد الأيام التي قضاها خمسون حدثا في أحد الدور الإصلاحية مستخدما بيانات الجدول رقم (٢٣ , ٤) بالفصل الثالث، والقاعدة العادية لإيجاده. ولتسهيل الحل نضع هذه البيانات في الصورة الموجودة في الجدول رقم (٤ , ٤) مع إضافة البيانات اللازمة لحساب المتوسط.

الجدول رقم (٤,٤). حساب المتوسط لفئات فترية.

الفئات	التكرار «ك»	مركز الفئة «س»	ك س
٢٤,٥ - ٢٩,٥	٦	٢٧	١٦٢
٢٩,٥ - ٣٤,٥	١٩	٣٢	٦٠٨
٣٤,٥ - ٣٩,٥	١٥	٣٧	٥٥٥
٣٩,٥ - ٤٤,٥	٧	٤٢	٢٩٤
٤٤,٥ - ٤٩,٥	٣	٤٧	١٤١
المجموع الكلي	٥٠ = مجـ ك	-	١٧٦٠ = مجـ ك س

من الجدول رقم (٤ , ٤) وباستخدام القاعدة:

$$\bar{س} = \frac{مجـ\ ك\ س}{مجـ\ ك}$$

وبالتعويض، فإن:

$$\bar{س} = \frac{١٧٦٠}{٥٠} = ٣٥,٢ \text{ يوما}$$

وهو متوسط عدد الأيام التي قضاها الخمسون حدثا بالدار الإصلاحية.

مثال (٤,٥)

أوجد الوسط الحسابي لعدد الأيام التي قضاها خمسون حدثا في الدور الإصلاحية مستخدما بيانات الجدول رقم (٢١ , ٣) بالفصل الثالث، وقاعدة الوسط الفرضي لإيجاده.

ولحل هذه المسألة نبدأ بوضع هذه البيانات مع إضافة الأعمدة الضرورية لحساب المتوسط باستخدام الوسط الفرضي. وسوف نختار مركز الفئة (٣٠-٣٤) وهو «٣٢» كوسط فرضي[٢] ومن ثم نكمل الأعمدة الضرورية لحساب المتوسط كما هو واضح من الجدول رقم (٥ , ٤).

الجدول رقم (٤,٥) . طريقة الوسط الفرضي لحساب المتوسط.

فئة اليوم بالسنة	مركز الفئة س	الانحراف ح = س - ف	التكرار ك	التكرار × الانحراف ك ح
٢٥-٢٩	٢٧	-٥	٦	-٣٠
٣٠-٣٤	٣٢	صفر	١٩	صفر
٣٥-٣٩	٣٧	٥	١٥	٧٥
٤٠-٤٤	٤٢	١٠	٧	٧٠
٤٥-٤٩	٤٧	١٥	٣	٤٥
المجموع الكلي	-	-	٥٠	١٦٠ = محـك ح

وبتطبيق القاعدة:

$$\bar{س} = ف + \frac{محـك ح}{محـك} \qquad (٦)$$

(٢) جرت العادة أن يختار مركز الفئة المنوالية، وهي الفئة التي يقابلها أعلى تكرار، كوسط فرضي وذلك لأن احتمال قربه من الوسط الحقيقي احتمال كبير.

وبالتعويض، فإن:

$$\bar{س} = ٣٢ + \frac{١٦٠}{٥٠}$$

$$= ٣٢ + ٣,٢$$

$$= ٣٥,٢ \text{ سنة}$$

وهي نفس الإجابة السابقة عند تطبيق القاعدة (٥).

٣,٢,٤ خواص الوسط الحسابي

للوسط الحسابي خاصيتان مهمتان يمكن تلخيصهما في القاعدتين:

(أ) $\text{مجـ} (س - \bar{س}) = \text{صفر}$ (٧)

(ب) $\text{مجـ} (س - \bar{س})^٢ = \text{أقل قيمة}$ (٨)

فالقاعدة (٧) تقول بالكلمات أن مجموع انحرافات القيم عن وسطها الحسابي يساوي صفرا، فيما تقول القاعدة (٨) أن مجموع تربيع انحرافات القيم عن وسطها الحسابي يكون دائما أقل من مجموع تربيع انحرافات تلك القيم عن أي قيمة تختلف عن الوسط الحسابي مهما كانت. ولإثبات صحة هاتين الخاصيتين دعنا نأخذ المثالين التاليين:

مثال (٤,٦)

أثبت أن مجموع انحراف القيم التالية عن وسطها الحسابي يساوي صفرا. والقيم هي: ١٠ ، ١٢ ، ٥ ، ٣ ، ٨.

الحل

نجد أولا الوسط الحسابي لهذه القيم بتطبيق القاعدة:

$$\bar{س} = \frac{\text{مجـ س}}{ن}$$

$$= \frac{١٠ + ١٢ + ٥ + ٣ + ٨}{٥} = \frac{٣٨}{٥} = ٧,٦$$

بعد ذلك نعوض في القانون (٧) فنحصل على :

مجـ (س - سَ) = (١٠ - ٧,٦) + (١٢ - ٧,٦) + (٥ - ٧,٦) + (٣ - ٧,٦)
+ (٨ - ٧,٦)
= ٢,٤ + ٤,٤ - ٢,٦ - ٤,٦ + ٠,٤
= ٧,٢ - ٧,٢ = صفر . . . وهو المطلوب .

مثال (٧، ٤)

باستخدام بيانات المثال (٦ ، ٤)، أثبت أن مجـ (س - سَ)^{٢} تساوي أقل قيمة .

ولحل هذه المسألة يتعين أن نجد أولا قيمة مجموع تربيع انحرافات القيم المدونة في المثال (٦ ، ٤) عن وسطها الحسابي ثم نجد ثانيا قيمة مجموع تربيع انحرافات نفس القيم عن أي قيمة أخرى تختلف عن الوسط الحسابي لهذه القيم . بعد ذلك نجري المقارنة بين الإجابتين حتى نتبين صحة ما تقول به القاعدة من أن الإجابة في الحالة الأولى يجب أن تكون أقل من الإجابة في الحالة الثانية على الترتيب . ويمكننا تكرار التجربة باتخاذ قيم مختلفة غير المتوسط الحسابي وإيجاد مجموع تربيع انحرافات قيم المثال عنها، وفي كل مرة نسجل الإجابة لنتأكد من أن مجموع تربيع انحرافات قيم المثال عن أي قيمة من هذه القيم التي تم اختيارها جزافا يكون أكبر من مجموع تربيع انحرافات قيم المثال عن وسطها الحسابي كما حسبناه أولا .

من المثال (٦ ، ٤)، فإن محصلة سلسلة (س - سَ) هي : (+ ٢,٤)، (+ ٤,٤)، (-٢,٦)، (- ٤,٦)، (+ ٠,٤). ومن هذه النتيجة فإن :

مجـ (س - سَ)^{٢} = (+ ٢,٤)^{٢} + (+ ٤,٤)^{٢} + (-٢,٦)^{٢} + (- ٤,٦)^{٢} + (+ ٠,٤)^{٢}
= ٥,٧٦ + ١٩,٣٦ + ٦,٧٦ + ٢١,١٦ + ٠,١٦ = ٥٣,٢ .

ولنأخذ الآن انحرافات قيم المثال (٦ ، ٤) عن أي قيمة أخرى غير متوسطها الحسابي

الذي هو ٦ , ٧ ، ولتكن مثلا العدد «٥». دعنا نعتبر هذه الـ «٥» كوسط فرضي ولنرمز إليه بالرمز $\bar{س}$. إذًا :

مجـ $(س - \bar{س})^2 = (١٠ - ٥)^2 + (١٢ - ٥)^2 + (٥ - ٥)^2 + (٣ - ٥)^2 + (٨ - ٥)^2$

= ٢٥ + ٤٩ + صفر + ٤ + ٩ = ٨٧ .

وهو أكبر من ٢ , ٥٣ وهذا هو المطلوب برهانه . ولسوف تتأكد للمبتدئ في الإحصاء أهمية هاتين الخاصيتين من خواص الوسط الحسابي إذا ما تقدم في العلوم الإحصائية .

٤ , ٢ , ٤ مزايا وعيوب الوسط الحسابي

٤ , ٢ , ٤ , ١ المزايا

١- هو أشهر المتوسطات جميعها وكثيرا ما لا يعرف من ليس لهم دراية بعلم الإحصاء متوسطا سواه .

٢- يمكن فهم فكرته دون مشقة لبساطتها الشديدة . كما أن حسابه سهل .

٣- يسهل التعامل معه جبريا مما أهله للدخول في كثير من العمليات الجبرية الخاصة بإيجاد مقاييس التشتت مثل الانحراف المعياري والتباين (انظر الفصل الخامس القادم) .

٤ , ٢ , ٤ , ٢ العيوب

١- لا يمكن الحصول عليه بطريقة الرسم .

٢- يتعذر حسابه في بعض الحالات مثل التوزيعات التكرارية التي تتضمن فئات مفتوحة [انظر الجدول رقم (٧ , ٣) مثلا] وليس بمستطاعنا قفلها .

٣- شديد التأثر بالقيم المتطرفة الكبيرة جدا أو الصغيرة جدا مما قد يجعله مضللا كمقياس للنزعة المركزية . مثال ذلك حساب متوسط دخل الفرد في دولة ثرية بينما تتركز الدخول العالية في يد فئة قليلة من المجتمع .

٤- غير مفيد في التوزيعات شديدة الالتواء وفي توزيعات البيانات الاسمية والرتبية.

٤.٣ المنوال

ويسمى أحيانا «الشائع» أو «النمط». والمنوال لأي مجموعة من القيم هو ببساطة «القيمة التي تتكرر أكثر من غيرها من القيم». ويميل الباحثون لاستخدام المنوال دون غيره من المتوسطات الأخرى في الحالات التي يودون فيها تعيين النمط العام أو الاتجاه العام للظاهرة أو للمتغير الذي هو قيد الدرس. ولذلك فهو يصلح لمعالجة المشكلات التي يتطلب السعي إلى حلها معرفة تركيز الظاهرة وموقعها وخاصة في الأمور التجارية مثل ضرورة الإلمام بموضات اللبس وأماكن شيوع هذه الموضات من أجل رواج تجارة الملابس.

٤.٣.١ إيجاد المنوال من البيانات غير المبوبة

مثال (٤.٨)

أوجد المنوال للقيم: ٤ ، ٢ ، ٥ ، ٣ ، ٤ ، ٥ ، ٢ ، ٣ ، ٣ ، ٤ .

الحل

المنوال = ٤ لأن العدد «٤» تكرر أكثر من غيره.

مثال (٤.٩)

أوجد المنوال للقيم: ٩ ، ١١ ، ٨ ، ٩ ، ١١ ، ١٥ ، ٨ ، ٩ ، ١٢ ، ١١ .

الحل

يوجد منوالان هما ٩ و ١١ لأن كل واحد من هذين العددين تكرر بنفس القدر، بينما تكرر كل واحد منهما أكثر من غيره من القيم. وتسمى مثل هذه المجموعة من

القيم مجموعة ذات منوالين. وبالطبع فإنه إذا وجد في التوزيع أكثر من منوال واحد فلا يمكننا استخدام المنوال كمقياس للنزعة المركزية.

مثال (٤,١٠)

أوجد المنوال للقيم: ١٥ ، ١١ ، ٩ ، ٨ ، ٧.

الحل

ليس لهذه القيم منوال لعدم تكرار أي قيمة منها.

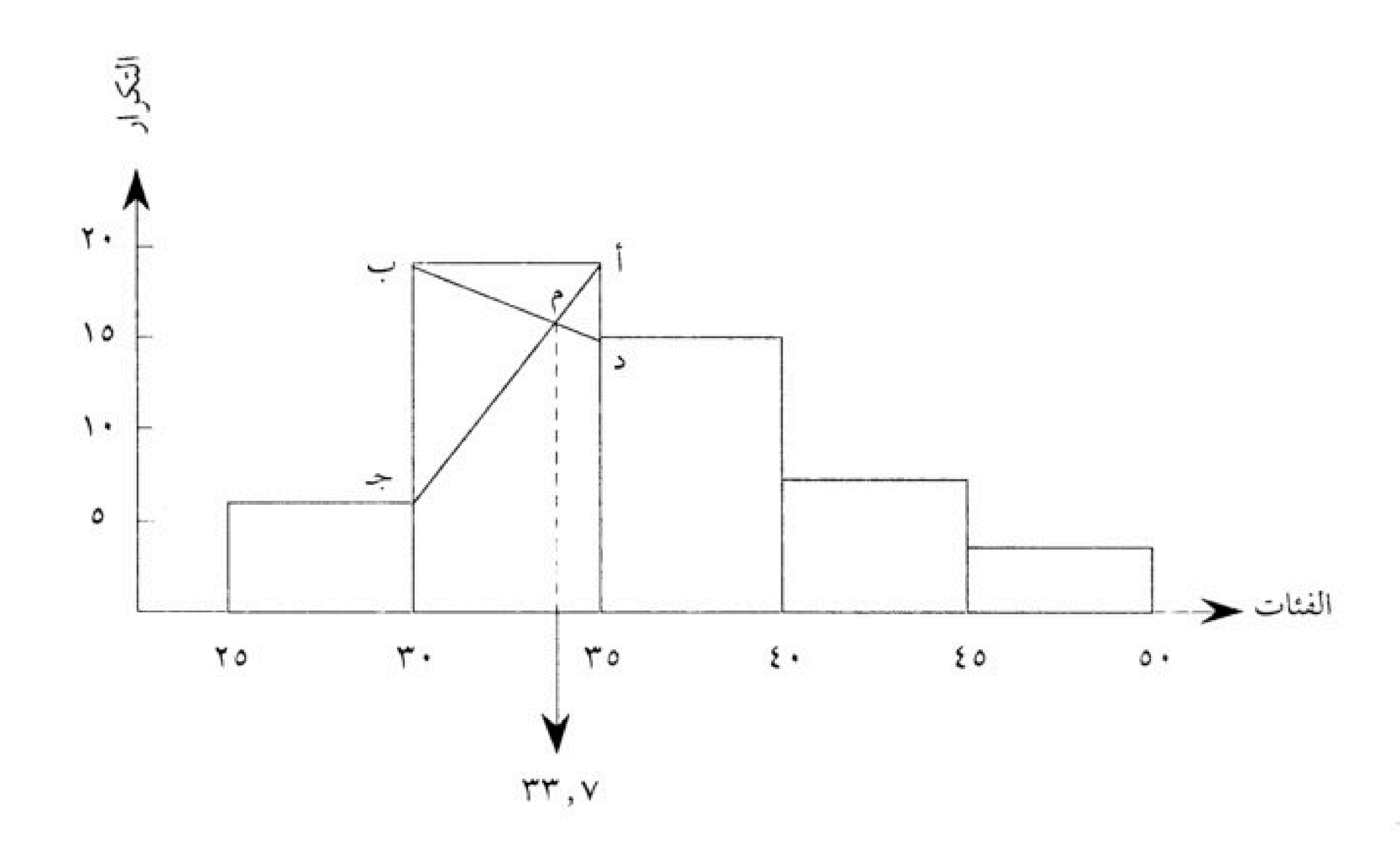
٤,٣,٢ إيجاد المنوال من البيانات المبوبة

وفق أي توزيع تكراري فإن الفئة التي يقابلها أكبر تكرار هي الفئة التي تحتوي على المنوال وتسمى الفئة المنوالية modal class . ولإيجاد المنوال باستخدام الرسم أو بالطريقة الحسابية فإننا لا نحتاج لأكثر من معرفة الفئة المنوالية والفئتين المجاورتين لها، أي الفئة السابقة لها والفئة اللاحقة لها.

٤,٣,٢,١ إيجاد المنوال بطريقة الرسم

دعنا نستخدم بيانات الجدول رقم (٢١ , ٣) لإيجاد المنوال عن طريق الرسم. وأول خطوة نقوم بها لتقدير المنوال عن طريق الرسم هي أن نرسم المدرج التكراري لبيانات الجدول رقم (٢١ , ٣) وهو كما يبدو في الشكل رقم (١ , ٤). ومتى ما مثلنا التوزيع التكراري بمدرج فإن الفئة المنوالية هي ذات أعلى مضلع (قد يكون المضلع مستطيلا أو مربعا أما في مثالنا هذا فهو مستطيل كما يرى في الشكل). ولإيجاد المنوال نقوم برسم ثلاثة مستطيلات تمثل تكرارات الفئة المنوالية (٣٠-٣٤) والفئة السابقة لها (٢٥-٢٩) ثم الفئة اللاحقة لها (٣٥-٣٩) [انظر الشكل]. بعد ذلك نقوم بتوصيل الركن الأيسر العلوي لمستطيل الفئة المنوالية (ب) بالركن الأيسر العلوي لمستطيل الفئة التالية لها (د) وكذلك نقوم بتوصيل الركن الأيمن العلوي لمستطيل الفئة المنوالية (أ) بالركن الأيمن

العلوي لمستطيل الفئة السابقة لها (جـ) ثم نسقط عمودا من نقطة التقاطع (م) إلى المحور الأفقي فتكون قيمة المنوال حيث يسقط العمود على المحور. وبقراءة عدد أيام هذه النقطة حسب مقياس الرسم الأفقي نجد أنها تساوي ٣٣,٧ يوما وهو منوال هذا التوزيع.



الشكل رقم (٤,١). كيفية إيجاد المنوال بطريقة الرسم.

٤,٣,٢,٢ إيجاد المنوال بطريقة الحساب

هناك قاعدتان يمكن استخدام أي منهما لإيجاد المنوال بطريقة الحساب إحداهما تسمى طريقة الرافعة لكنج King's method، وتقرأ كما يلي:

$$\text{المنوال} = \text{أ} + \frac{\text{ك}_{٢}}{\text{ك}_{١} + \text{ك}_{٢}} \times \text{ط} \qquad (٩)$$

حيث:

أ = بداية الفئة المنوالية.

$\text{ك}_{١}$ = تكرار الفئة السابقة للفئة المنوالية.

$ك_٢$ = تكرار الفئة اللاحقة للفئة المنوالية .

ط = طول الفئة المنوالية .

وبتطبيق هذه القاعدة على المثال السابق، أي على بيانات الجدول (٢١ , ٣)، وبالتعويض فإننا سوف نحصل على :

$$\text{المنوال} = ٣٠ + \frac{١٥}{٦ + ١٥} \times ٥$$

$$= ٣٠ + \frac{١٥}{٢١} \times ٥ = ٣٠ + ٣,٥٧ = ٣٣,٦ \text{ يوما تقريبا .}$$

والطريقة الأخرى هي طريقة الفروق لكارل بيرسون Pearson's method ولقد صممت هذه الطريقة لتلافي عيب طريقة كنج الرئيسي والمتمثل في أنها لا تدخل في الاعتبار أكبر تكرار في التوزيع، وهو تكرار الفئة المنوالية، عند التقدير . وتقرأ طريقة «بيرسون» كما يلي :

$$\text{المنوال} = أ + \frac{ك - ك_١}{(ك - ك_١) + (ك - ك_٢)} \times ط \qquad (١٠)$$

حيث :

أ = بداية الفئة المنوالية .

ك = تكرار الفئة المنوالية .

$ك_١$ = تكرار الفئة السابقة للفئة المنوالية .

$ك_٢$ = تكرار الفئة اللاحقة للفئة المنوالية .

ط = طول الفئة المنوالية .

وبتطبيق القاعدة (١٠) على المثال السابق نفسه، أي على بيانات الجدول رقم (٢١ , ٣)، وبالتعويض فإننا سوف نحصل على :

$$\text{المنوال} = ٣٠ + \frac{١٩ - ٦}{(١٩ - ٦) + (١٩ - ١٥)} \times ٥$$

$$= ٣٠ + \frac{١٣}{١٣ + ٤} \times ٥$$

$$= ٣٠ + \frac{١٣}{١٧} \times ٥ = ٣٠ + ٣,٨ = ٣٣,٨ \text{ يوما تقريبا .}$$

٤,٤ الوسيط The Median

إذا تم ترتيب مجموعة من القيم ترتيبا تصاعديا أو تنازليا فإن الوسيط هو «القيمة التي تقع عند منتصف المجموعة بالضبط» . وبمعنى آخر فإن الوسيط هو «القيمة التي تقسم مثل هذه المجموعة إلى نصفين بحيث يكون عدد القيم الأصغر منها مساويا تماما لعدد القيم الأكبر منها» .

٤,٤,١ إيجاد الوسيط من البيانات غير المبوبة

يكون الوسيط في حالة البيانات غير المبوبة هو القيمة ذات المرتبة $\left(\frac{ن + ١}{٢}\right)$ حينما يكون عدد القيم (ن) فرديا ، بينما يكون هو الوسط الحسابي للقيمتين اللتين لهما الرتبتين $\left(\frac{ن}{٢}\right)$ و $\left(\frac{ن}{٢} + ١\right)$ إذا كان عدد القيم (ن) زوجيا . ولنأخذ الآن مثالا تطبيقيا لإيجاد الوسيط عندما يكون عدد القيم (ن) فرديا ومثالا آخر عندما يكون عدد القيم (ن) زوجيا .

مثال (٤,١١)

أوجد الوسيط لمجموعة القيم التالية :

٩ ، ٤ ، ٥ ، ٩ ، ٦ ، ٥ ، ٩ ، ٧ ، ١١

لحل هذه المسألة نقوم أولا بترتيب القيم ، وليكن ترتيبا تصاعديا حيث يقرأ كما يلي :

٤ ، ٥ ، ٥ ، ٦ ، ٧ ، ٩ ، ٩ ، ٩ ، ١١ وحيث إن ن = ٩ وهو عدد فردي ، لذلك فإن رتبة الوسيط لهذه المجموعة توجد باستخدام المعادلة :

$$\frac{ن + ١}{٢} \qquad (١١)$$

وبالتعويض فإن ترتيب الوسيط = $\frac{٩ + ١}{٢}$ = ٥ . وبما أن القيمة التي لها الرتبة ٥ في المجموعة المرتبة هي ٧ ، يكون الوسيط = ٧ .

مثال (٤,١٢)

أوجد الوسيط لمجموعة القيم التالية :

١١ ، ١٤ ، ٢٠ ، ٩ ، ٧ ، ١٣ ، ٧ ، ١٧ وحيث إن ن = ٨ وهو عدد زوجي ، فإننا نطبق القاعدة التالية ذات الشقين لإيجاد رُتْبَتَيْ القيمتين اللتين سوف نحسب متوسطهما ليمثل الوسيط لمجموعة القيم أعلاه . والقاعدة ذات الشقين هي :

$$\left\{ \begin{array}{l} \text{رتبة القيمة الأولى} = \frac{ن}{٢} \\ \text{رتبة القيمة الثانية} = \frac{ن}{٢} + ١ \end{array} \right. \qquad (١٢)$$

وبالتعويض في القاعدة (١٢) عن ن بالقيمة ٨ نجد أن :

رتبة القيمة الأولى = $\frac{٨}{٢}$ = ٤ وهذا العدد يشير إلى رتبة إحدى القيمتين المعنيتين .

رتبة القيمة الثانية = $\frac{٨}{٢}$ + ١ = ٤ + ١ = ٥ وهذا العدد يشير إلى القيمة الأخرى المعنية .

وإذا نظرنا إلى مجموعة القيم في المثال (١٢ , ٤) فإن ترتيبها تصاعديا يكون كما يلي : ٧ ، ٧ ، ٩ ، ١١ ، ١٣ ، ١٤ ، ١٧ ، ٢٠ . وبالتالي تكون القيمة ذات الترتيب

٤ هي ١١ ، والقيمة ذات الترتيب ٥ هي ١٣ . وعلى ذلك يكون الوسيط لهذه المجموعة من القيم هو :

$$\text{الوسيط} = \frac{١١ + ١٣}{٢} = \frac{٢٤}{٢} = ١٢ .$$

٤,٤,٢ إيجاد الوسيط من البيانات المبوبة

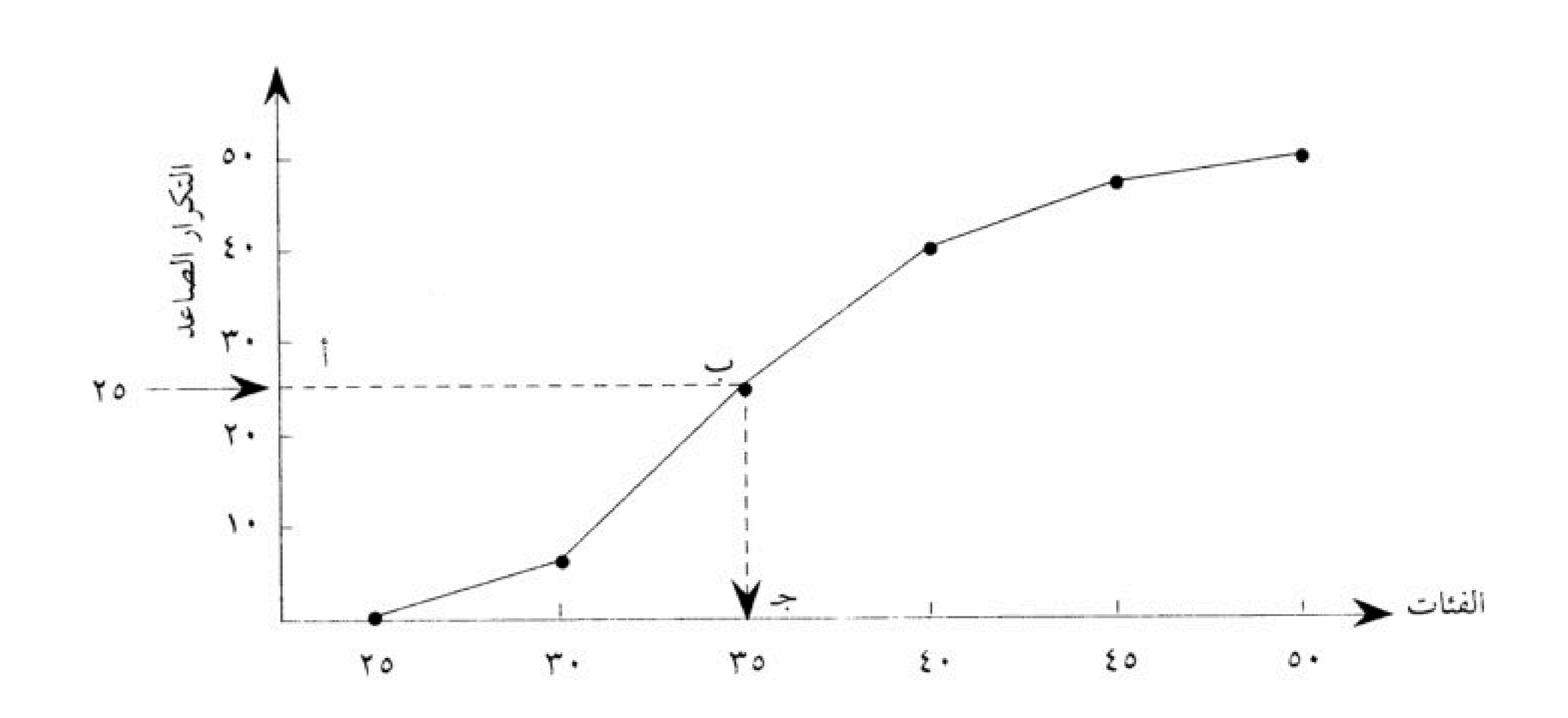
٤,٤,٢,١ طريقة الرسم

لإيجاد الوسيط من البيانات المبوبة باستخدام طريقة الرسم نبدأ بحساب التكرار المتجمع الصاعد ثم نرسم المنحنى التكراري الذي يقابله . دعنا نستخدم بيانات الجدول رقم (٦ , ٤) ليكون هو مثالنا في حساب الوسيط من البيانات المبوبة . الشكل رقم (٢ , ٤) يوضح رسم المنحنى التكراري المتجمع الصاعد وطريقة تعيين نقطة الوسيط . ولإيجاد الوسيط من الشكل رقم (٢ , ٤) نبدأ بتعيين نقطة ترتيبه (أ) على المحور الرأسي بأخذ نصف عدد المفردات كما يلي :

الجدول رقم (٦ ,٤). الجدول التكراري المتجمع الصاعد لخمسين حدثاً.

الفئة	التكرار	التكرار المتجمع الصاعد
٢٥-٢٩	٦	٦
٣٠-٣٤	١٩	٢٥
٣٥-٣٩	١٥	٤٠
٤٠-٤٤	٧	٤٧
٤٥-٤٩	٣	٥٠
المجموع	٥٠	-

$\frac{ن}{٢} = \frac{محـك}{٢} = \frac{٥٠}{٢} = ٢٥$. ثم نرسم من النقطة «أ» خطا يوازي المحور الأفقي ليقطع المنحنى المتجمع الصاعد عند النقطة «ب» . ومن النقطة «ب» نرسم خطا يوازي المحور الرأسي للتكرارات المتجمعة الصاعدة ليلاقي المحور الأفقي عند النقطة «جـ» التي تساوي قراءتها الأفقية الوسيط وهو ٣٥ يوما (انظر الشكل) .



الشكل رقم (٢ ,٤). كيفية إيجاد الوسيط بطريقة الرسم.

٢ ,٢ ,٤ ,٤ طريقة الحساب

لإيجاد الوسيط من البيانات المبوبة بطريقة الحساب نستخدم القاعدة :

$$الوسيط = ل_و + \frac{\frac{ن}{٢} - (محـك)_و}{ك_و} \times ط_و \qquad (١٣)$$

حيث :

$ل_و$ = الحد الأدنى للفئة الوسيطية (أي التي يقع فيها الوسيط) .

ن = عدد مفردات البيانات (أي المجموع الكلي للتكرارات محـك) .

(مجـ ك)$_و$ = مجموع التكرارات لجميع الفئات قبل الفئة الوسيطية.
ك$_و$ = تكرار الفئة الوسيطية.
ط$_و$ = طول الفئة الوسيطية.

وكمثال تطبيقي لإيجاد الوسيط بطريقة الحساب دعنا نأخذ مرة أخرى مثال التوزيع التكراري كما هو موضح في الجدول رقم (٧ , ٤).

الجدول رقم (٧.٤). إيجاد الوسيط بطريقة الحساب.

الفئة	التكرار	التكرار المتجمع الصاعد	عدد التكرارات التي تقل أيامها عن:
٢٥-٢٩	٦	٦	٣٠
٣٠-٣٤	١٩	٢٥	٣٥
٣٥-٣٩	١٥	٤٠	٤٠
٤٠-٤٤	٧	٤٧	٤٥
٤٥-٤٩	٣	٥٠	٥٠
المجموع	٥٠	-	-

وباستخدام بيانات الجدول رقم (٧ , ٤)، وبالتعويض في القاعدة (١٣) يكون الوسيط هو:

$$\text{الوسيط} = ٣٠ + \frac{\frac{٥٠}{٢} - ٦}{١٩} \times ٥$$

$$= ٣٠ + \frac{٢٥ - ٦}{١٩} \times ٥$$

$$= ٣٠ + \frac{١٩}{١٩} \times ٥$$

= ٣٠ + ٥ = ٣٥ يوما . وهي نفس الإجابة عند إيجاده بالرسم .

٤,٥ أسئلة

١- أوجد الوسط الحسابي والوسيط والمنوال لمجموعة القيم الآتية :
١٠ ، ١٦ ، ١٨ ، ١٣ ، ١٦

٢- اتخذ الرقم ٨ كوسط فرضي لحساب الوسط الحسابي للقيم المدونة في السؤال (١) .

٣- تتبع مستشفى حكومي تطور نمو عدد من الأطفال يتبعون لخمسة مراكز صحية يشرف عليها ، فأعطى كل مجموعة منهم مؤشرا للصحة العامة لفترة زمنية حددها المستشفى ذاته فحصل على التوزيع الآتي :

المركز	عدد الأطفال	مؤشر الصحة
الأول	٨	٧,٥
الثاني	٧	٩
الثالث	١٢	٤,٥
الرابع	٩	١٣
الخامس	١٩	٥,٥

المطلوب :
أوجد بطريقة الوسط الحسابي المرجح الوسط الحسابي لمؤشر الصحة العام لجميع الأطفال .

٤- ضع الأرقام التالية في جدول تكراري مناسب بحيث تستطيع استخدام القاعدة: $\bar{س} = \frac{مجـ ك س}{مجـ ك}$ لإيجاد وسطها الحسابي:

٨، ٦، ٤، ٧، ٦، ٥، ٦، ٣، ٩، ٧، ٦، ٥

٩، ٤، ٥، ٦، ٥، ٩، ٣، ٦، ٥.

٥- أوجد المنوال والوسيط لأرقام المسألة (٤).

٦- من الجدول التالي الذي يوضح توزيع عدد من الطلاب حسب الدرجات التي تحصلوا عليها في اختبار لمادة الإحصاء، أوجد الوسط الحسابي والوسيط والمنوال لدرجات هؤلاء الطلاب.

فئة الدرجات	٢٠-٢٩	٣٠-٣٩	٤٠-٤٩	٥٠-٥٩	٦٠-٦٩	٧٠-٧٩	٨٠-٨٩
عدد الطلاب	٤	٦	١٢	١٤	٩	٣	٢

٤,٦ اصطلاحات ينبغي تذكرها

- النزعة المركزية.
- المتوسطات.
- الوسط الحسابي المرجح.
- الوسط الفرضي.
- انحرافات القيم.
- طريقة كنج لحساب الوسيط.
- طريقة بيرسون لحساب الوسيط.
- مربعات النهاية الصغرى.

الفصل الخامس

مقاييس التشتت
Measures of Dispersion

٥,١ تمهيد

بالإضافة إلى المتوسطات أو مقاييس النزعة المركزية هناك عدة مقاييس إحصائية جميعها تعين على إبراز المعالم الرئيسة لمجموعة البيانات التي يحصل عليها الباحث، منها مقاييس التشتت measures of dispersion بشقيها التشتت المطلق والنسبي absolute and relative dispersion، ومقاييس الالتواء measures of skewness، ومقاييس التفرطح measures of kurtosis . ولأغراض هذا المؤلف سوف نتناول في فصلنا هذا مقاييس التشتت فقط، المطلق منها والنسبي، إذْ إن هذا الفصل يعتبر مكملا للفصل الرابع الذي سبقه عن المتوسطات لأن هذه الأخيرة وحدها لا تكفي لإعطاء صورة متكاملة لمجموعة من القيم. ذلك أنه قد يحدث أن نجد مجموعتين من القيم لهما نفس المتوسط والوسيط لكنهما تختلفان اختلافا كبيرا في المدى، مثلا، وهو مقياس للتشتت، مما يتعذر معه تكوين فكرة واضحة عن درجة التجانس بينهما. ولتعضيد هذا المذهب بمثال تطبيقي خذ مثلا أوزان مجموعة تتكون من خمسة طلاب كما يلي:

٤٠ ، ٤٤ ، ٤٢ ، ٤٦ ، ٤٨ كجم.

وأخرى أوزانها:

٢٠ ، ٦٤ ، ٤٢ ، ٣٦ ، ٥٨ كجم.

فمتوسط وزن الطالب بالمجموعة الأولى = $\frac{٤٠ + ٤٤ + ٤٢ + ٤٦ + ٤٨}{٥}$ = ٤٤ كجم.

بينما المدى the range لهذه المجموعة هو ٤٨ - ٤٠ = ٨ كجم فقط. أما متوسط وزن الطالب بالمجموعة الثانية فهو $\frac{٢٠ + ٦٤ + ٤٢ + ٣٦ + ٥٨}{٥}$ = ٤٤ كجم أيضا، بينما المدى لهذه المجموعة الثانية هو ٦٤ - ٤٠ = ٢٤ كجم وهو ثلاثة أضعاف مدى المجموعة الأولى مما يعكس اختلافا واضحا بين المجموعتين فيما يتعلق بدرجة انتشار قيم كل منهما حول الوسط نفسه. ولأسباب كهذه، ولأجل الحصول على صورة أمثل عن أي مجموعة من القيم لا بد لنا من معرفة أحد متوسطاتها وكذلك أحد مقاييس انتشارها جنبا إلى جنب حتى نكون في وضع أدعى للحكم على درجة التجانس بين تلك القيم.

إن مقاييس التشتت المطلق ومقاييس التشتت النسبي هي التي تفي بغرض إكمال وصف صورة توزيع القيم عندما تؤخذ جنبا إلى جنب مع مقاييس النزعة المركزية. فبينما يعطينا مقياس التشتت المطلق الانتشار أو التشتت في القيم كعدد من الوحدات التي استخدمت في قياس القيم المعطاة، يعطينا مقياس التشتت النسبي التشتت أو الانتشار كنسبة فقط. ويتميز المقياس الأخير - أي مقياس التشتت النسبي - على الأول في أنه يمكننا من مقارنة المجموعات ذات المتوسطات المختلفة بينما يفقد مقياس التشتت المطلق هذه الخاصية. وتتكون أهم مقاييس التشتت المطلق من المدى، ونصف المدى الربيعي semi-interquartile range والانحراف المتوسط mean deviation، والانحراف المعياري standard deviation والتباين variance. كما أن أهم مقاييس التشتت النسبي تتمثل في الدرجة المعيارية standardized (or Z) score ومعامل الاختلاف coefficient of variation. وسوف نتناول كيفية حساب مقاييس التشتت المطلق والنسبي المذكورة فيما يلي من عرض.

٥,٢ المـدى

المدى هو الفرق بين أكبر قيمة وأصغر قيمة في أي مجموعة من القيم، كما سبقت الإشارة إلى ذلك، عندما تكون البيانات غير مبوبة. أما عندما تكون البيانات

مبوبة فالأمر يختلف اختلافا طفيفا تمليه طبيعة تنظيم البيانات في جدول وسوف نتطرق إلى ذلك في حينه. وسوف نتناول الآن حساب المدى في حالتي البيانات غير المبوبة والبيانات المبوبة كلا على حدة.

٥,٢,١ حساب المدى من البيانات غير المبوبة

مثال (٥,١)

أوجد المدى للقيم: ٤٠ ، ٤٤ ، ٤٢ ، ٤٦ ، ٤٨ .

الحـــل

أكبر قيمة = ٤٨

أقل قيمة = ٤٠

إذًا المدى = ٤٨ - ٤٠ = ٨ .

ويعطى المدى في بعض الأحيان بذكر أقل وأكبر قيمة، وعلى هذا يمكن أن نعبر عن مدى هذه المجموعة من القيم بأنه من ٤٠ إلى ٤٨ .

٥,٢,٢ حساب المدى من البيانات المبوبة

يمكن تعريف المدى للبيانات المبوبة ذات الفئات التي لها حدود دنيا وأخرى عليا بأنه:

المدى = الحد الأعلى لأكبر فئة - الحد الأدنى لأصغر فئة

وإذا أردنا تطبيق التعريف السابق على مثالنا التقليدي لبيانات الجدول رقم (٣,٢١) لحساب المدى لعدد الأيام التي قضاها خمسون حدثا في إحدى الدور الإصلاحية فيمكننا حسابه كالآتي:

الحد الأعلى لأكبر فئة = ٤٩

الحد الأدنى لأصغر فئة = ٢٥

وبذلك يكون

المدى = ٤٩ – ٢٥ = ٢٤ يوما .

٥,٢,٣ مزايا المدى وعيوبه

يمتاز المدى بأنه :

١ – سهل الحساب بحيث لا يتطلب معرفة معمقة بالرياضيات أو الإحصاء .

٢ – يعطي فكرة سريعة وموجزة عن طبيعة البيانات .

أما عيوبه فيمكن تلخيصها في أنه :

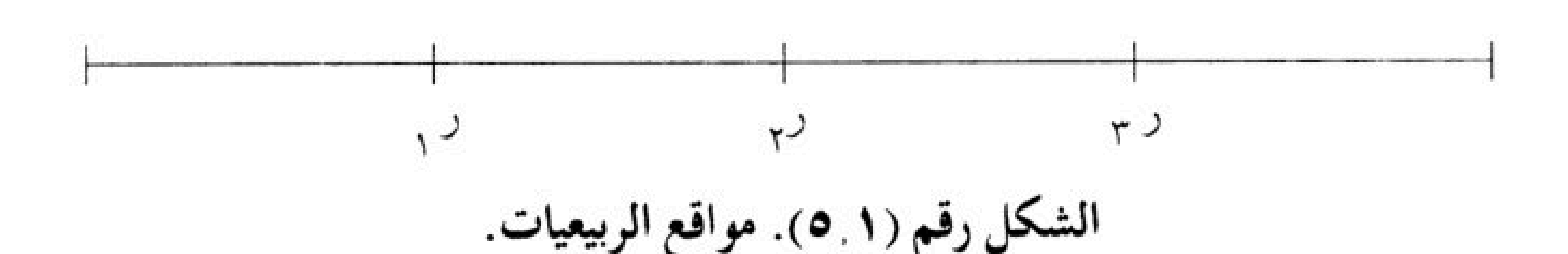
١ – مقياس غير دقيق ويتأثر كثيرا بالقيم المتطرفة أو الشاذة .

٢ – لا يمكن استخدامه في حالة البيانات الوصفية .

٥,٣ نصف المدى الربيعي

يرتبط نصف المدى الربيعي – كمقياس للتشتت – ارتباطا وثيقا بالربيع الأول (أو الربيع الأدنى) والربيع الثالث (أو الربيع الأعلى) وهما شبيهان بالوسيط إلا أننا لم نتطرق إليهما ضمن هذه عند تناولها في الفصل السابق . إلا أن التطرق إلى نصف المدى الربيعي يحتم علينا تناولهما الآن لتتضح فكرته بصورة أدق .

إذا كان لدينا خط متصل وقمنا بتقسيمه إلى أربعة أجزاء متساوية الطول ؛ فإن كل جزء من هذه الأجزاء يمثل ربع الخط المتصل الذي يمثل بالضرورة مجموعة بياناتنا الكمية . نرمز للربع الأول من هذه البيانات – والتي يتحتم أن يتم ترتيبها تصاعديا أو تنازليا ونبدأ تسمية الربيعيات بدءًا من جهة القيم الدنيا – بالرمز $ر_١$ وللربع الثاني بالرمز $ر_٢$ ، وللربع الثالث $ر_٣$ كما هو واضح في الشكل رقم (١ , ٥) :



الشكل رقم (٥,١). مواقع الربيعيات.

من الشكل رقم (١ , ٥) فإن:

المدى الربيعي $= ر_٣ - ر_١$ (١٤)

نصف المدى الربيعي $= \frac{ر_٣ - ر_١}{٢}$ (١٥)

ويجدر بالذكر أن $ر_٢$ ما هو إلا الوسيط لهذه البيانات.

الآن وقد اتضحت فكرة الربيعيات نتدرج إلى كيفية حساب نصف المدى الربيعي.

١,٣,٥ حساب نصف المدى الربيعي من البيانات غير المبوبة

لحساب نصف المدى الربيعي في حالة البيانات غير المبوبة يلزم أولا ترتيب القيم المعطاة ترتيبا تصاعديا أو تنازليا ثم الاضطلاع بإيجاد قيم $ر_١$ و $ر_٣$ لنؤسس بذلك لإيجاد نصف المدى الربيعي الذي يعنى بتطبيق القاعدة (١٥). وأول الخطى في هذا الطريق هو إيجاد رتبة أو ترتيب الربيع الأول ($ر_١$) باستخدام القاعدة:

رتبة $ر_١ = \frac{ن}{٤}$ (١٦)

ثم نجد رتبة الربيع الثالث ($ر_٣$) بتطبيق القاعدة:

رتبة $ر_٣ = \frac{٣ن}{٤}$ (١٧)

حيث ترمز (ن) في كلا القاعدتين إلى عدد القيم المعطاة. وفي حالة عدم قبول «ن»

القسمة على ٤ ، نأخذ القيمتين اللتين يقع بينهما العدد الكسري ونحسب وسطهما الحسابي لنحصل بعدئذ على قيمتي $ر_{١}$ و $ر_{٣}$ كما سوف يتضح فيما يأتي من أمثلة .

مثال (٥,٢)

احسب نصف المدى الربيعي للقيم التالية :

١٢ ، ١١ ، ٩ ، ٨ ، ١٧ ، ٥ ، ٣ ، ١٣ .

الحـل

نقوم أولا بترتيب مجموعة القيم ترتيبا تصاعديا فيكون لدينا سلسلة القيم :

القيمة	٣	٥	٨	٩	١١	١٢	١٣	١٧
الرتبة	١	٢	٣	٤	٥	٦	٧	٨

رتبة $ر_{١}$ $= \frac{ن}{٤} = \frac{٨}{٤} = ٢$ ، إذًا قيمة $ر_{١} = ٥$

رتبة $ر_{٣}$ $= \frac{٣ن}{٤} = \frac{٣ \times ٨}{٤} = ٦$ ، إذًا قيمة $ر_{٣} = ١٢$

نصف المدى الربيعي $(ر) = \frac{ر_{٣} - ر_{١}}{٢} = \frac{١٢ - ٥}{٢} = ٣,٥$.

مثال (٥,٣)

احسب نصف المدى الربيعي للبيانات التالية :

١٣ ، ٣ ، ٥ ، ١٧ ، ٨ ، ٩ ، ١١ ، ١٢ ، ٧ .

الحـــل

نرتب القيم ترتيبا تصاعديا كما يلي :

القيمة	٣	٥	٧	٨	٩	١١	١٢	١٣	١٧
الرتبة	١	٢	٣	٤	٥	٦	٧	٨	٩

قبل البدء في التعويض لحل المسألة، نلاحظ أن عدد القيم « ن » = ٩ وهو عدد لا يقبل القسمة على ٤ . وعليه يكون هناك معالجة خاصة كما سبق وأن نوّه إلى ذلك سابقا .

إذًا، رتبة $ر_{١}$ = $\frac{ن}{٤}$ = $\frac{٩}{٤}$ = ٢,٢٥ . نأخذ القيمتين اللتين لهما الرتبتان : ٢ و٣ وهما القيمتان ٥ و ٧ كما يظهر في الجدول أعلاه، ثم نوجد متوسط هاتين القيمتين ليمثل لنا قيمة $ر_{١}$ ، أي أن $ر_{١}$ = $\frac{٥ + ٧}{٢}$ = $\frac{١٢}{٢}$ = ٦ . ونحسب الآن قيمة $ر_{٣}$:

$ر_{٣}$ = $\frac{٣ن}{٤}$ = $\frac{٣ \times ٩}{٤}$ = $\frac{٢٧}{٤}$ = ٦,٧٥ ؛ فنأخذ القيمتين ذواتي الرتبتين ٦ و٧ وهما القيمتان ١١ و١٢ . ثم بعد ذلك نوجد متوسطهما ليمثل قيمة $ر_{٣}$، أي أن :

$ر_{٣}$ = $\frac{١١ + ١٢}{٢}$ = $\frac{٢٣}{٢}$ = ١١,٥ . الآن نستطيع إيجاد نصف المدى الربيعي ($ر$) =

$\frac{ر_{٣} - ر_{١}}{٢}$ = $\frac{١١,٥ - ٦}{٢}$ = ٢,٧٥ . ويلاحظ أن قيمة المدى الربيعي (وليس نصف المدى الربيعي) توجد ببساطة بإزالة العدد «٢» عن مقام القاعدة (١٥) .

٢,٣,٥ حساب نصف المدى الربيعي من البيانات المبوبة

إن أيسر وأدق طريقة لإيجاد نصف المدى الربيعي من البيانات المبوبة هي طريقة إيجاده باستخدام الرسم البياني، وبالتحديد الاستعانة برسم المنحنى المتجمع الصاعد للبيانات المبوبة. وسوف نستخدم هنا أيضا بيانات الجدول رقم (٣ , ٢١) كمثال نطبق عليه عملية إيجاد (ر) بالاستعانة برسم منحنى البيانات لتكراراته المتجمعة الصاعدة، ويمكن أن نكتب هذه البيانات كما في الجدول رقم (٥ , ١).

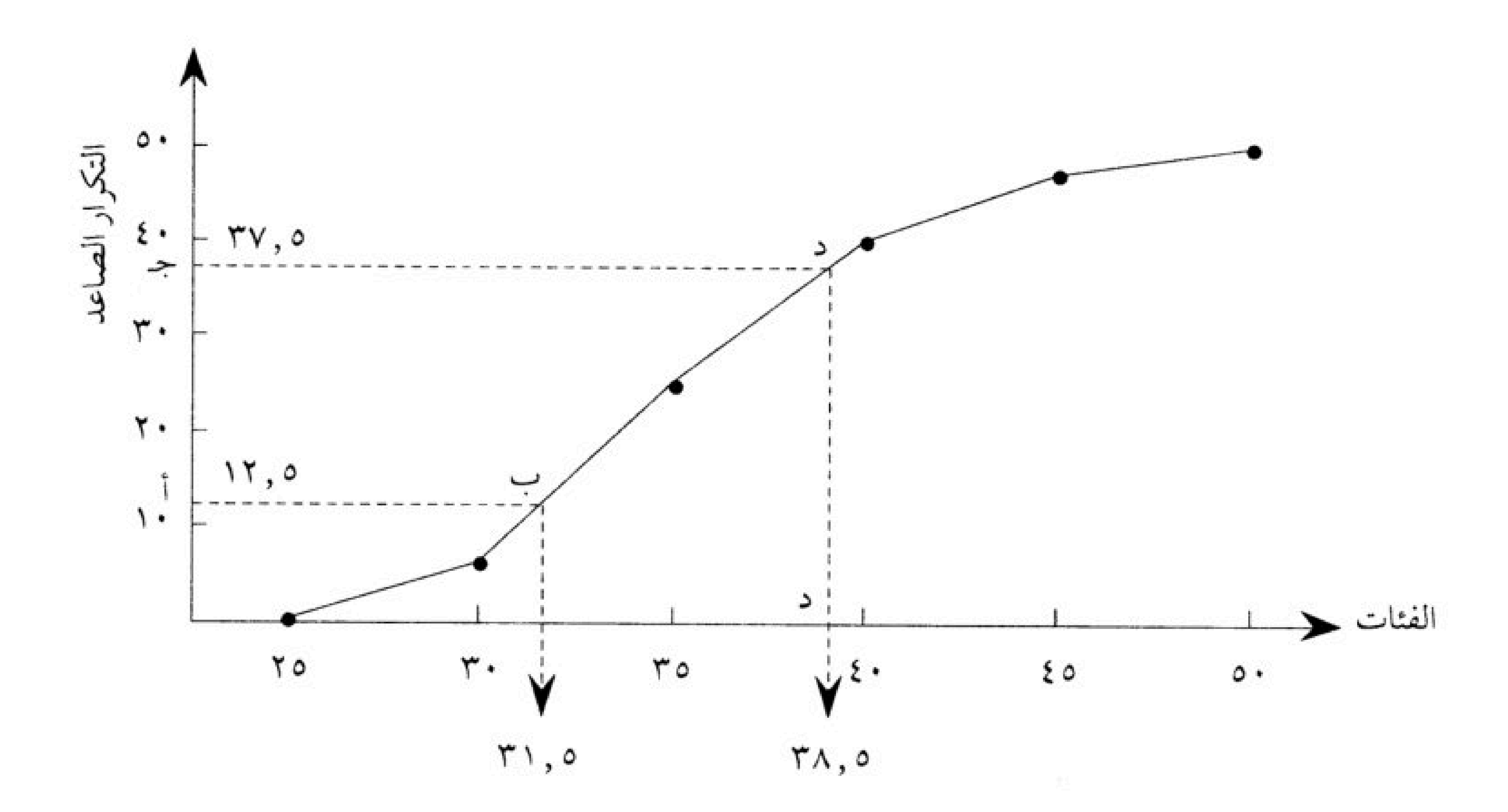
الجدول رقم (٥,١) . استخدام المنحنى المتجمع الصاعد لإيجاد (ر).

الفئات	٢٥-٢٩	٣٠-٣٤	٣٥-٣٩	٤٠-٤٤	٤٥-٤٩
التكرار	٦	١٩	١٥	٧	٣
التكرار الصاعد	٦	٢٥	٤٠	٤٧	٥٠

من بيانات الجدول رقم (٥ , ١)، نوجد رتبة قيمة ر_1، وهي $= \frac{\text{ن}}{٤} = \frac{٥٠}{٤} =$ ١٢,٥ . ثم نوجد رتبة قيمة ر_3 وهي $\frac{٣\text{ن}}{٤} = \frac{٣ \times ٥٠}{٤} =$ ٣٧,٥ . ثم نرسم من النقطة (أ) في الشكل رقم (٥ , ٢) والتي تشير إلى القيمة ($\text{ر}_1 =$ ١٢,٥) خطا مستقيما يوازي المحور الأفقي للشكل (محور الفئات) ونسير به إلى أن يتقاطع مع المنحنى المتجمع الصاعد في النقطة (ب). بعد ذلك نسقط عمودا من النقطة ب على محور الفئات (انظر الخط المتقطع) ونقرأ النقطة التي يسقط عليها العمود على هذا المحور ويتبين أنها تساوي ٣١,٥ . وهذه هي بالضبط قيمة ر_1 . ونجري نفس العملية

لإيجاد قيمة $ر_{٣}$ برسم خط مواز للمحور الأفقي يبدأ من النقطة (جـ) على محور التكرارات المتجمعة ليلاقي المنحنى الصاعد عند النقطة (د) حيث نسقط منها عمودا على محور الفئات ونقرأ نقطة تقاطع العمود مع هذا المحور والتي يتبين أنها تساوي ٣٨,٥ وهي بالضبط قيمة $ر_{٣}$. بعد ذلك نطبق القاعدة (١٥) لنجد أن (ر) لهذا

$$\text{التوزيع} = \frac{ر_{٣} - ر_{١}}{٢} \text{ ، أي } = \frac{٣٨,٥ - ٣١,٥}{٢} = ٣,٥ \text{ يوما .}$$



الشكل رقم (٥,٢). إيجاد $ر_{١}$ و $ر_{٣}$ من بيانات الجدول رقم (٥,١) بمنحنى صاعد.

٥,٣,٣ مزايا وعيوب نصف المدى الربيعي

تتلخص مزايا نصف المدى الربيعي في أنه :

١- قادر على التخلص من القيم المتطرفة سواء كانت تشذ عن غيرها من القيم بكبرها أو بصغرها .

٢- يمكن حسابه في حالة التوزيعات التكرارية التي بها فئات مفتوحة .

أما عيوب نصف المدى الربيعي فيمكن حصرها في :

١- لا يأخذ جميع قيم التوزيع في الاعتبار عند حسابه .

٢- يصعب التعامل معه جبريا .

٥,٤ الانحراف المتوسط (أو متوسط الانحرافات)

نعرف من طرح سابق أن مجموع الانحرافات الحقيقية للقيم عن وسطها الحسابي يساوي دائما صفرا ، أي أن مجـ (س - سَ) = صفرا . ولما كان من الممكن دراسة درجة التشتت لأي مجموعة من القيم بمدى تباعد أو انتشار مفرداتها عن المتوسط العام لهذه القيم يمكننا الاستفادة من انحرافات القيم «المطلقة» أي قيم انحرافات المجموعة مع إهمال إشاراتها (الإشارة السالبة تعامل وكأنها موجبة) لإيجاد مؤشر للتشتت يطلق عليه اسم متوسط الانحرافات ، ويعرف بأنه مجموع الانحرافات المطلقة مقسوما على عددها .

٥,٤,١ إيجاد الانحراف المتوسط من البيانات غير المبوبة

إذا كانت لدينا مجموعة تتكون من «ن» من القيم ورمزنا إليها بـ : $س_١$ ، $س_٢$ ،

$س_٣$ ، . . . ، $س_ن$ فإن الانحراف المتوسط لهذه المجموعة يوجد بتطبيق القاعدة :

$$\text{الانحراف المتوسط} = \frac{\text{مجـ} \,|س - \bar{س}|}{ن} \qquad (١٦)$$

حيث :

$\bar{س}$ = الوسط الحسابي للقيم .

$|س - \bar{س}|$ = الانحراف المطلق لأي قيمة عن الوسط الحسابي .

$مجـ |س - \bar{س}|$ = مجموع الانحرافات المطلقة .

$ن$ = عدد القيم (أو عدد انحرافات القيم جميعها) .

مثال (٤,٥)

احسب الانحراف المتوسط للقيم التالية :

٥ ، ٧ ، ٩ ، ١٢ ، ١٧ .

الحل

لتسهيل الحل نقوم بتكون الجدول التالي :

						المجموع
القراءات (س)	٥	٧	٩	١٢	١٧	٥٠
الانحرافات $|س - \bar{س}|$	٥	٣	١	٢	٧	١٨

من الجدول السابق ، فإن $\bar{س} = \frac{مجـ س}{ن} = \frac{٥٠}{٥} = ١٠$. وبطرح ١٠ من كل

قيمة من القيم المثبتة في الصف الأول تحصلنا على القيم المثبتة في الصف الثاني والتي تمثل الانحرافات المطلقة عن المتوسط ١٠ ، ومجموعها هو ١٨ . إذًا :

$$\frac{مجـ |س - \bar{س}|}{ن} = \frac{١٨}{٥} = ٣,٦ .$$

وهذه هي قيمة الانحراف المتوسط للبيانات السابقة .

٢,٤,٥ إيجاد الانحراف المتوسط من البيانات المبوبة

لإيجاد الانحراف المتوسط من البيانات المبوبة ذات الفئات تعدل القاعدة (١٦) بإدخال متغير التكرارات عليها حيث تكون قاعدة إيجاد الانحراف المتوسط من هذا النوع من البيانات كما يلي:

$$\text{الانحراف المتوسط} = \frac{ك_{١}\left|س_{١} - \bar{س}\right| + ك_{٢}\left|س_{٢} - \bar{س}\right| + \ldots + ك_{م}\left|س_{م} - \bar{س}\right|}{ك_{١} + ك_{٢} + \ldots + ك_{م}}$$

$$\text{إذًا، الانحراف المتوسط} = \frac{\text{مجـ } ك\left|س - \bar{س}\right|}{\text{مجـ } ك} \qquad (١٧)$$

حيث «ك» يشير إلى تكرار الفئة.

وفي حالة الفئات التي لها حدود دنيا وعليا، فإن «س» تشير إلى مراكز الفئات.

مثال (٥,٥)

نرجع إلى مثالنا التقليدي وهو بيانات الجدول رقم (٣,٢١) فنجد لها انحرافها المتوسط. ولتسهيل خطوات الحل نضع بيانات ذلك الجدول في الصورة الموجودة في الجدول رقم (٥,٢).

من الجدول رقم (٥,٢) نجد $\bar{س} = \frac{\text{مجـ } ك\, س}{\text{مجـ } ك} = \frac{١٧٦٠}{٥٠} = ٣٥,٢$. ولقد استخدمنا الوسط الحسابي بقيمته ٣٥,٢ لإيجاد الانحرافات المطلقة في العمود الرابع من على يمين الجدول. وبتطبيق القاعدة (١٧) فإن:

$$\text{متوسط الانحرافات} = \frac{\text{مجـ } ك\left|س - \bar{س}\right|}{\text{مجـ } ك} = \frac{٢٢٠}{٥٠} = ٤,٤ \text{ يوما.}$$

الجدول رقم (٥,٢). إيجاد الانحراف المتوسط من البيانات المبوبة.

فئات الأيام	التكرارات ك	مراكز الفئات س	|س - سَ|	ك|س - سَ|	ك س
٢٥-٢٩	٦	٢٧	٨,٢	٤٩,٢	١٦٢
٣٠-٣٤	١٩	٣٢	٣,٢	٦٠,٨	٦٠٨
٣٥-٣٩	١٥	٣٧	١,٨	٢٧,٠	٥٥٥
٤٠-٤٤	٧	٤٢	٦,٨	٤٧,٦	٢٩٤
٤٥-٤٩	٣	٤٧	١١,٨	٣٥,٤	١٤١
المجموع	٥٠	-	-	٢٢٠,٠	١٧٦٠

٥,٥ الانحراف المعياري والتباين

إذا كانت لدينا مجموعة من القيم تتألف من « ن » قيمة، ورمزنا لهذه القيم بـ:

$س_١$، $س_٢$، $س_٣$، . . . ، $س_ن$ فإن الانحراف المعياري لهذه المجموعة هو:

$$ع = \sqrt{\frac{مجـ (س - \bar{س})^٢}{ن}} \qquad (١٨)$$

حيث:

$(س - \bar{س})^٢$ = مربع انحراف أية قيمة عن الوسط الحسابي للقيم.

مجـ ك $(س - \bar{س})^٢$ = مجموع مربعات انحرافات القيم عن وسطها الحسابي.

$$\frac{\text{مجـ}\,(\text{س} - \bar{\text{س}})^{2}}{\text{ن}}$$ = متوسط مربعات انحرافات القيم عن وسطها الحسابي .

ع = الانحراف المعياري .

وعليه، فإن الانحراف المعياري يمكن تعريفه بأنه «الجذر التربيعي لمتوسط مربعات انحرافات القيم عن وسطها الحسابي» .

أما التباين لمجموع «ن» من القيم «رمزنا لمفرداتها بنفس الرموز أعلاه» فيمكن كتابة معادلته كالتالي :

$$\text{التباين} = \text{ع}^{2} = \frac{\text{مجـ}\,(\text{س} - \bar{\text{س}})^{2}}{\text{ن}} \qquad (١٩)$$

وهذا يعني أن التباين يمكن أن يعرف بأنه «متوسط مربعات انحرافات القيم عن وسطها الحسابي» .

وفي حالة ما إذا كانت بياناتنا مبوبة، فإن معادلة الانحراف المعياري أعلاها يمكن أن تعدل لتقرأ :

$$\text{ع} = \sqrt{\frac{\text{مجـ ك}\,(\text{س} - \bar{\text{س}})^{2}}{\text{مجـ ك}}} \qquad (٢٠)$$

حيث :

ك = تكرار الفئة .

س = الفئة أو مركز الفئة في حالة ما إذا كان لهذه حدان .

على أن هناك عدة معادلات بديلة لإيجاد الانحراف المعياري سواء كان يستخرج من البيانات المبوبة أو غير المبوبة وسوف نعرض لذكر بعضها مع اتصال الطرح . أما إيجاد التباين (ع^{2}) من البيانات المبوبة فيوجد – عند استخدام المعادلة (٢٠) لإيجاد الانحراف

المعياري – فقط بإزالة علامة الجذر التربيعي عن المعادلة (٢٠)، لتكون معادلة التباين بذلك :

$$\text{التباين} = ع^{٢} = \frac{\text{مجـ ك} (س - \bar{س})^{٢}}{\text{مجـ ك}} \quad (٢١)$$

١,٥,٥ إيجاد الانحراف المعياري من البيانات غير المبوبة

مثال (٥,٦)

أوجد الانحراف المعياري «ع» للقيم الآتية : ٢ ، ٥ ، ٣ ، ٧ ، ٨ .

الحـل

لتسهيل عملية الحساب نضع البيانات أعلاه كما في الجدول رقم (٥,٣) .

الجدول رقم (٥,٣). إيجاد الانحراف المعياري من بيانات غير مبوبة.

س	$(س - \bar{س})$	$(س - \bar{س})^{٢}$
٢	-٣	٩
٥	صفر	صفر
٣	-٢	٤
٧	٢	٤
٨	٣	٩
٢٥ = مجـ س	-	٢٦ = مجـ $(س - \bar{س})^{٢}$

من البيانات المعطاة فإن «ن» وهو عدد القيم يساوي ٥ $\bar{س} = \frac{مجـ س}{ن} = \frac{٢٥}{٥} = ٥$.

إذن نجد المتوسط . ثم نجد انحراف أية قيمة من قيم س بالعمود الأول من جهة اليمين عن ٥ ، ونثبت سلسلة هذه الانحرافات في العمود الثاني من جهة اليمين . أما في العمود الثالث والأخير فقد وضعنا قيم مربعات تلك الانحرافات . ولإيجاد الانحراف المعياري «ع»، ما علينا إلا أن نطبق القاعدة (١٨) فنكتب :

$$ع = \sqrt{\frac{مجـ (س - \bar{س})^{٢}}{ن}}$$

$$= \sqrt{\frac{٢٦}{٥}}$$

$$= \sqrt{٥,٢} = ٢,٣ \text{ تقريبا}.$$

وجدير بالذكر ، أن التباين $ع^{٢}$ = ٥ , ٢ ، وذلك بإزالة علامة الجذر التربيعي عن المعادلة (١٨) كما سبق وأن ذكرنا .

٢ ,٥ ,٥ إيجاد الانحراف المعياري من البيانات المبوبة

مثال (٧, ٥)

أوجد الانحراف المعياري «ع» لبيانات الجدول رقم (٢١ , ٣) .

ولحل هذه المسألة ننظم بيانات الجدول المعني في الصورة الموضحة في الجدول رقم (٤ , ٥)، وذلك لتسهيل عملية الحل .

الجدول رقم (٤.٥). إيجاد الانحراف المعياري من البيانات المبوبة.

فئة الأيام	مركز الفئة س	(س - سَ) = (س - ٣٥,٢)	(س - سَ)² = (س - ٣٥,٢)²	ك	ك (س - سَ)²
٢٥-٢٩	٢٧	-٨,٢	٦٧,٢٤	٦	٤٠٣,٤٤
٣٠-٣٤	٣٢	-٣,٢	١٠,٢٤	١٩	١٩٤,٥٦
٣٥-٣٩	٣٧	١,٨	٣,٢٤	١٥	٤٨,٦٠
٤٠-٤٤	٤٢	٦,٨	٤٦,٢٤	٧	٣٢٣,٦٨
٤٥-٤٩	٤٧	١١,٨	١٣٩,٢٤	٣	٤١٧,٧٢
المجموع الكلي	-	-		٥٠ = مجـك	١٣٨٨,٠ = مجـك (س - سَ)²

لقد عرفنا من معلوماتنا السابقة أن الوسط الحسابي لهذه البيانات يساوي ٣٥,٢ يوما، ولذلك لم نتجشم عملية إجراء حسابه من جديد حيث افترضنا أنها معلومة معطاة سلفا. ومن نتائج معالجة البيانات في الجدول أعلاه، ما علينا إلا تعويض الأرقام في القاعدة (٢٠) للحصول على الانحراف المعياري «ع» كما يلي:

$$ع = \sqrt{\frac{\text{مجـك (س - سَ)}^2}{\text{مجـك}}}$$

$$= \sqrt{\frac{١٣٨٨}{٥٠}}$$

$$= \sqrt{٢٧,٧٦} = ٥,٢٧ = ٥,٣$$ يوما تقريبا.

وإذا طلب منا إيجاد التباين لهذا التوزيع التكراري فما علينا بالطبع إلا إزالة علامة الجذر التربيعي عن القيمة ٢٧,٧٦ لتكون هي بالتالي قيمة التباين «ع٢» لهذا التوزيع.

٣,٥,٥ معادلات أخرى مختصرة لإيجاد الانحراف المعياري من نوعي البيانات

(أ) $ع = \frac{١}{ن}\sqrt{ن\ مجـ\ س^{٢} - (مجـ\ س)^{٢}}$ بيانات غير مبوبة (٢٢)

(ب) $ع = \frac{١}{ن}\sqrt{ن\ مجـ\ ك\ س^{٢} - مجـ\ (ك\ س)^{٢}}$ بيانات مبوبة (٢٣)

(جـ) $ع = \sqrt{\frac{مجـ\ س^{٢}}{ن} - \left(\frac{مجـ\ س}{ن}\right)^{٢}}$ بيانات غير مبوبة (٢٤)

(د) $ع = \sqrt{\frac{مجـ\ ك\ س^{٢}}{مجـ\ ك} - \left(\frac{مجـ\ ك\ س}{مجـ\ ك}\right)^{٢}}$ بيانات مبوبة (٢٥)

(هـ) $ع = \sqrt{\frac{مجـ\ ح^{٢}}{ن} - \left(\frac{مجـ\ ح}{ن}\right)^{٢}}$ بيانات غير مبوبة (٢٦)

حيث ح تساوي انحراف أية قيمة (س) عن الوسط الحسابي للقيم.

(و) $ع = \sqrt{\frac{مجـ\ ك\ ح^{٢}}{مجـ\ ك} - \left(\frac{مجـ\ ك\ ح}{مجـ\ ك}\right)^{٢}}$ بيانات مبوبة (٢٧)

وحيث «ح» تساوي انحراف أية قيمة عن الوسط الحسابي للقيم.

٥,٦ مقاييس التشتت النسبي

خصصت مقاييس التشتت النسبي للتعبير عن التشتت أو التباين كنسبة مئوية إلى مقياس النزعة المركزية أو المتوسط الذي قيست منه انحرافات القيم . ونعلم من خبراتنا السابقة أن المتوسط الذي تقاس منه انحرافات القيم قد يكون الوسط الحسابي أو الوسيط أو المنوال ، وإن كان أول هذه المقاييس - أي الوسط الحسابي - هو الأكثر استخداما بينها . وتستخدم مقاييس التشتت النسبي عادة لمقارنة التباين في توزيع ما بالتباين في توزيع آخر . ونتناول فيما يلي مقياسين من مقاييس التشتت النسبي ، هما الدرجة المعيارية ومعامل الاختلاف .

٥,٦,١ إيجاد الدرجة المعيارية ومعامل الاختلاف

لكل من الدرجة المعيارية (د) ومعامل الاختلاف (م) معادلة بسيطة لإيجاده .

ففي الحالة الأولى : $$\text{د} = \frac{\text{س} - \bar{\text{س}}}{\text{ع}} \qquad (٢٨)$$

وفي الحالة الثانية : $$\text{م} = \frac{\text{ع}}{\bar{\text{س}}} \times ١٠٠ \qquad (٢٩)$$

مثال (٥,٨)

حصل طالب على ٦٠ درجة (الدرجة الكاملة ١٠٠) في اختبار لمادة الإحصاء وعلى ٧٠ درجة (الدرجة الكاملة ١٠٠) في اختبار لمادة الاجتماع . وكان الوسط الحسابي لدرجات الطلاب في كل من المادتين ، على الترتيب ، ٥٤ و٦٥ درجة . كما أن الانحراف المعياري لهما على الترتيب ، ٢ و٥ , ٢ درجة .

(أ) مستخدما مفهوم الدرجة المعيارية (د) بيّن في أي المادتين يكون مستوى الطالب أفضل؟

(ب) مستخدما مفهوم معامل الاختلاف (م) بيّن في أي المادتين كان تحصيل الطلاب أكثر تباينا؟

الحل

لحل هذه المسألة يحسن أن نضع البيانات المعطاة في جدول حتى يسهل عملية معالجتها . الجدول رقم (٥ , ٥) ينظم المعطيات .

الجدول رقم (٥,٥). إيجاد الدرجة المعيارية ومعامل الاختلاف.

المعطيات المادة	عدد الدرجات س	الوسط الحسابي $\bar{س}$	الانحراف المعياري ع
الإحصاء	٦٠	٥٤	٢
الاجتماع	٧٠	٦٥	٢,٥

(١) إجابة الجزء (أ) من المثال:

القاعدة العامة : $د = \frac{س - \bar{س}}{ع}$

الدرجة المعيارية لمادة الإحصاء : $د = \frac{٦٠ - ٥٤}{٢} = \frac{٦}{٢} = ٣$

الدرجة المعيارية لمادة الاجتماع : $د = \frac{٧٠ - ٦٥}{٢,٥} = \frac{٥}{٢,٥} = ٢$

إذن :

مستوى الطالب في مادة الإحصاء أفضل منه في مادة الاجتماع حيث إن درجته في الإحصاء تبعد عن المتوسط لهذه المادة بـ٣ وحدات من الانحراف المعياري ، بينما تبعد

درجته في الاجتماع عن المتوسط لهذه المادة بوحدتين (٢) فقط من الانحراف المعياري.

ملحوظة مهمة

يمكن تعريف الدرجة المعيارية رياضيا بأنها عدد الوحدات من الانحراف المعياري التي تبعد بها الدرجة الخام عن الوسط الحسابي (أو أي متوسط آخر).

(٢) إجابة الجزء (ب) من المثال:

القاعدة العامة: $م = \frac{ع}{\bar{س}} \times ١٠٠$

معامل الاختلاف لدرجات الطلاب في مادة الإحصاء: $م = \frac{٢}{٥٤} \times ١٠٠ = ٣,٧٪$

معامل الاختلاف لدرجات الطلاب في مادة الاجتماع: $م = \frac{٢,٥}{٦٥} \times ١٠٠ = ٣,٨٪$

ونخلص مما سبق من نتيجتي معاملي الاختلاف لتوزيع درجات الطلاب في المادتين أن تحصيلهم كان أكثر تباينا (أو أقل تجانسا) في مادة الاجتماع (م = ٣,٨٪) عنه في مادة الإحصاء (م = ٣,٧٪).

٥,٧ أسئلة

١- فيما يلي ثلاث مجموعات من القيم يُطلب ترتيب قيم كل مجموعة ترتيبا تصاعديا ومن ثم الحصول على الربيع الأول والثاني والثالث لكل مجموعة، ثم إيجاد نصف المدى الربيعي لكل مجموعة أيضا.

(أ) ٣، ٦، ١١، ١٠، ٥، ٨، ٤، ٧، ٩.

(ب) ٤٠، ٣٤، ٤٣، ٤٠، ٣١، ٢٨، ٢٩.

(جـ) ٤٩، ٤٠، ٢٥، ٤٥، ٥٣، ٥٧، ٣٤، ٤١، ٤٣، ٤٣.

٢- أوجد الوسط الحسابي ثم الانحراف المتوسط لكل مجموعة من المجموعات الثلاث في السؤال رقم (١).

٣- أوجد الانحراف المعياري لكل مجموعة من الثلاث مجموعات المدونة قيمها في السؤال رقم (١) بطريقتين (أي معادلتين) مختلفتين.

٤- استخدم طريقة الوسط الفرضي لحساب الانحراف المعياري لكل مجموعة من المجموعات الثلاث في السؤال رقم (١).

٥- فيما يلي توزيع ٥٠ طالبا حسب فئات الدرجات التي تحصلوا عليها في مادة الإحصاء.

الفئة	١٠-	٢٠-	٣٠-	٤٠-	٥٠-	٦٠-	٧٠-	٨٠-	٩٠-١٠٠	المجموع
التكرار	صفر	٤	٦	١٢	١٤	٩	٣	٢	صفر	٥٠

(أ) استخدم طريقة الوسط الفرضي لحساب الانحراف المعياري للتوزيع.

(ب) احسب الوسط الحسابي للتوزيع.

(جـ) أوجد نصف المدى الربيعي للتوزيع.

٦- حصل طالب على ٨٦ درجة في مادة الاجتماع التي كان الوسط الحسابي لدرجاتها ٧٨ وانحرافها المعياري ١٢، وعلى ٩٢ درجة في مادة الاقتصاد التي كان متوسط درجاتها ٨٤ وانحرافها المعياري ١٨.

(أ) في أي المادتين كان أداء الطالب أفضل؟

(ب) في أي المادتين كان تباين الأداء عموما أكبر؟

٥,٨ اصطلاحات ينبغي تذكرها

- المتوسطات.
- مقاييس النزعة المركزية.
- مقاييس التشتت المطلق.
- مقاييس التشتت النسبي.
- الانحراف المتوسط/ متوسط الانحرافات.
- الدرجة المعيارية.
- معامل الاختلاف.
- الانحرافات المطلقة.
- الوسط الفرضي.
- نصف المدى الربيعي:
- الانحراف المعياري/ التباين.

الفصل السادس

الارتباط Correlation

يكون المبتدئ في علم الإحصاء الاجتماعي الذي درس هذا المؤلّف من بدايته إلى نهاية الفصل الخامس قد ألفَ من أدوات المعالجة الإحصائية للبيانات الخام ما يؤهله لوصف أي ظاهرة اجتماعية بأساليب تُحرّض فهم الشخص العادي على إدراك السمات البارزة والخصائص المميزة لتلك الظاهرة. إذ إنه علاوة على استطاعته اختزال البيانات الخام الضخمة والمبعثرة في جدول تكراري صغير يفسح المجال لتصور ذهني مريح لمختلف أوجه تلك الظاهرة بأرقام منطقية يمكن تحويلها إلى نسب ومعدلات وقيم تناسب أدق وضعا، يستطيع الدارس كذلك أن يحول تلك التكرارات البسيطة إلى رسومات بيانية تكاد تنطق وصفا للظاهرة لمن لا يستطيع قراءة الأرقام، أو لا يستسيغها بقدر يوازي استساغته أو حتى حبه المدرجات والمضلعات والمنحنيات ومختلف الأشكال الهندسية التي يمكن لتلك الأرقام أن تبدعها. ليس هذا وحسب، وإنما باستطاعة الدارس المعني أيضا أن يصف مدى نزوع قيم أي ظاهرة نحو قيمة وسط تلخّص درجة تمركزها، ومدى نزوع تلك القيم نحو قيمة بعينها تصف مدى انتشارها حول القيمة الوسط، وكذلك يستطيع أن يصف ما إذا كان التجانس (أو التباين) بين قيم ظاهرة ما أكثر أو أقل منه بين قيم ظاهرة أخرى. كل هذا مطلوب، بل وأساسي في مفاصل البحث الاجتماعي غير أن متطلبات البحث الاجتماعي الحديث كانت ولا تزال تواقة دوما إلى الانطلاق به نحو آفاق أرحب بكثير مما تعلمناه حتى الآن. فالظاهرة التي ندرسها قد تكون مشكلة نسعى لحلها. ولا يكفي لحل المشكلة أن نكون مهرة في وصفها بل يتطلب الأمر إدراك مسبباتها. وهذا بدوره يتطلب دراسة العلاقات بينها وبين مشكلات أخرى قد تكون مرتبطة بها حتى نتمكن من فهم طبيعة المشكلة بشكل أدق قبل أن نكون مؤهلين علميا للإقبال على وضع الحلول المناسبة لها. فمثلا،

إذا وجدنا أن متوسط الدخل الشهري لمجموعة من الأشخاص متدنٍ كثيرا مقارنة بمستوى الأسعار السائد في المنطقة التي تسكنها هذه المجموعة، فهل يُعزى ذلك إلى انخفاض مستوى تعليم هؤلاء الأشخاص؟ أم إلى انكماش الطلب على الخدمات والسلع التي ينتجونها؟ أم إلى عجز الحد الأدنى للأجور التي يتقاضونها عن اللحاق بمستوى معدلات التضخم بالمنطقة؟ أم إلى غير ذلك من مختلف الأسباب والمسببات التي يمكن أن تتضافر لتؤدي إلى متوسط الدخل الشهري الملاحظ لهذه المجموعة؟ والذي لم يتعد جهدنا في تعلم مبادئ الإحصاء الاجتماعي حتى الآن - لم يتعد إتاحة إمكانية وصفه بأنه متدن دونما عبور نحو تفسير أسباب هذا التدني توطئة لوصف جرعات العلاج المناسبة. وإن إحدى أهم طرق التحليل الإحصائي التي تساعد على التمهيد لولوج باب تفسير الظواهر أو المتغيرات الاجتماعية التي هي قيد الدرس تتمثل في ما يطلق عليه معاملات الارتباط correlation coefficients والتي سوف نتناول أبرزها بالشرح فيما يلي من طرح.

٦,١ معاملات الارتباط

قد يزعم أحد الباحثين الاجتماعيين أن هناك علاقة بين عزوف الشباب الذكور عن الزواج وغلاء المهور، أو أن هناك علاقة بين أسلوب تربية الأبناء وانحراف الأحداث منهم، أو أن هناك علاقة بين مستوى دخل الفرد ومستوى تعليمه، أو أن هناك علاقة بين الطبقة الاجتماعية والتحصيل الدراسي، أو أن هناك علاقة بين أي متغير اجتماعي مماثل وأكثر من متغير واحد. وربما لا يكتفي مثل هذا الباحث بمثل زعم كهذا بل من الممكن أن يبذل مزيدا من الاجتهاد في ملاحظاته فيزعم أنه كلما ازداد وقع متغير من متغيراته المزعومة تلك ازداد بالتالي وقع المتغير الآخر الذي يزعم أن له علاقة به، أو كلما نقص وقع ذلك المتغير زاد وقع المتغير الآخر المصاحب له، أو العكس. وربما يكون في وارد ملاحظاته أيضا أن العلاقة أو الارتباط بين متغيرين من تلك المتغيرات متلازمة تماما وأقوى بصورة تفوق كثيرا قوة العلاقة بين متغيرين آخرين. ولكن كيف يتسنى لمثل هذا الباحث أن يترجم ملاحظاته تلك إلى مقاييس أو إلى أرقام عددية يقنع بها المراقب بصحة ما يزعم من ملاحظات؟ هذا سؤال يتكفل بالإجابة عليه واضعوا علم الإحصاء حين اهتدوا إلى اكتشاف مقياس كمي للعلاقة بين متغيرين

لا يضطلع بتقدير كمي لقوة العلاقة بين متغيرين فحسب وإنما يتعدى ذلك إلى تحديد اتجاه تلك العلاقة هل هي طردية (موجبة) أم عكسية (سالبة) بين المتغيرين، ويتبين هذا فقط من نوع الإشارة التي تسبق القيمة التقديرية لقوة العلاقة، إذ تُحسب، هل هي موجبة أم سالبة. ويطلق على مثل هذا المقياس الكمي للعلاقة بين متغيرين اسم «معامل الارتباط». وأبرز معاملات الارتباط نوعان: أحدهما يسمى معامل ارتباط بيرسون «ر» Pearson's r ويتم استخدامه لوصف العلاقة بين متغيرين مقاسين على مستوى القياس الفاصل interval أو النسبي ratio scale على أن تكون العلاقة بينهما خطية(*) وتوزيعهما طبيعيا، والآخر يسمى معامل ارتباط سبيرمان للرتب (رر) Spearman's r ويشترط لتطبيقه قياس المتغيرين على مستوى القياس الرتبي مع استيفاء شرطي خطية العلاقة بينهما وطبيعية توزيع قراءات المتغيرين كما هو الحال مع المتغيرين الفاصلين. وهناك مقاييس أخرى لكشف العلاقات بين المتغيرات التي لا يلزم أن يكون توزيعها طبيعيا ويمكن أن تكون مقاسة على ميزان القياس الاسمي كذلك وسوف نتعرض إلى أهم هذه المقاييس بالدرس والنقاش في حينه.

وإذا أردنا قياس العلاقة بين عدد سنوات التعليم والدخل الشهري لمجموعة أشخاص مثلا؛ نبدأ أولا بتأمل شكل العلاقة بين هذين المتغيرين بـ «شكل الانتشار» scattergram ولدرجات المجموعة في المتغيرين حيث يتم وضعها داخل محورين متقاطعين عموديا وبمقياسي رسم رأسي (ص) وأفقي (س) كما هو واضح في الشكل رقم (١ , ٦). وشكل الانتشار عبارة عن عدد من النقاط تساوي في مجموعها عدد مفردات الدراسة. وكل نقطة من هذه النقاط تمثل قراءتين لمفردة واحدة: إحداهما تمثل درجة المفردة في متغير الدخل الشهري في مثالنا بالشكل رقم (١ , ٦) والأخرى تمثل درجتها في متغير عدد سنوات الدراسة بنفس الشكل. فإذا كانت العلاقة خطية بين المتغيرين كما هو الحال في مثال الشكل رقم (١ , ٦) أو الشكل رقم (٢ , ٦) فإن المسار العام لنقاط الانتشار يأخذ شكلا يمكن التعبير عنه بخط مستقيم يشطر هذه النقاط إلى نصفين متكافئين تقريبا وعندئذ يمكننا القول إن العلاقة بين المتغيرين خطية. وإذا تأكدنا من أن العلاقة بينهما خطية بفحص المسار العام للنقاط يمكن بعد ذلك الشروع في حساب معامل الارتباط «ر» لقياس قوة واتجاه العلاقة بين المتغيرين. وعلى ذكر قوة العلاقة بين المتغيرين نقرر هنا بأنها تقاس بقيمة «ر» التي تتراوح بين +١ صحيح و-١ صحيح، ويعبر عنها كالتالي:

-١ ≤ ر ≤ ١

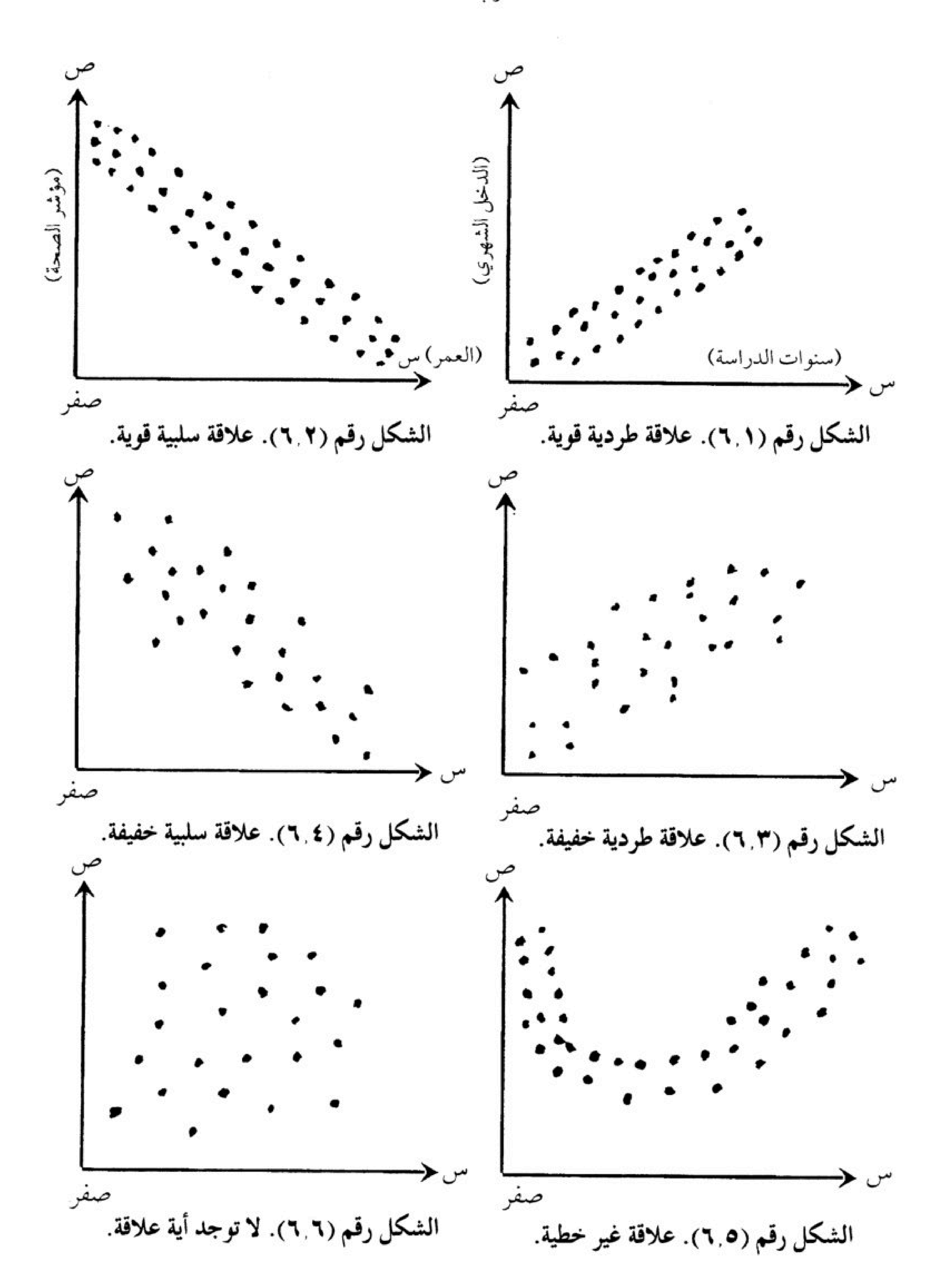
فإذا كانت العلاقة تامة أو كاملة بين المتغيرين فإن مقدار «ر» = + ١ صحيح، أو = -١ صحيح والإشارة التي تسبق العدد «١» تدل على أن العلاقة طردية (أو إيجابية) بين المتغيرين إذا كانت (+) وعلى أنها عكسية (أو سالبة) بينهما إذا كانت (-) وعندما تكون نقاط الانتشار أكثر قربا من بعضها في مسارها العام فإن هذا يكون مؤشرا على أن العلاقة قوية بين المتغيرين، وإذا كانت هذه النقاط تميل إلى التباعد عن بعضها في المسار العام لها فإن هذا يكون مؤشرا على أن العلاقة تميل إلى الضعف بين المتغيرين كما في الشكلين رقمي (٣ , ٦)، (٤ , ٦). ويمكن للدارس أن يتخيل عددا لا حصر له من مدى قوة أو ضعف الارتباط بين المتغيرات الذي يمكن ترجمته إلى مختلف أشكال الانتشار المعبرة عنه. وهناك حالات لا تكون فيها العلاقة بين المتغيرين خطية كما في الشكل رقم (٥ , ٦) أو لا تكون هناك علاقة أصلا بين المتغيرين كما يتضح من الشكل رقم (٦ , ٦) وفي هذه الحالة تكون قيمة «ر» = صفر.

أما فيما يتعلق باتجاه العلاقة بين المتغيرين فإذا كان المسار العام للنقاط يمكن التعبير عنه بخط مستقيم يماثل قطرا ينصف الزاوية القائمة (٩٠°) التي يصنعها تعامد الإحداثيين س، ص كما يعكس ذلك الشكل رقم (١ , ٦) فإن العلاقة تكون موجبة (أي طردية)، أما إذا كان من الممكن التعبير عن هذاك المسار العام بخط مستقيم يمثل قاطعا من أعلى الإحداثي الصادي إلى النهاية اليمنى للإحداثي السيني كما يوضح ذلك الشكل رقم (٢ , ٦) فإن العلاقة تكون في الاتجاه السالب (أي عكسية).

وجدير بالذكر أنه ليس هناك قاعدة محددة يمكن اتباعها لوصف العلاقة بين متغيرين بأنها علاقة قوية أو ضعيفة سوى أنه معلوم بأن العلاقة تكون أقوى كلما كان مقدار «ر» أقرب إلى الـ «١» الصحيح (سواء كان موجبا أو سالبا) وتكون أضعف كلما كان مقدار «ر» أبعد منه. وعلى أية حال يمكن الاهتداء ببيانات الجدول رقم (١ , ٦)[١] الذي أقر ذوو الخبرة في أبحاث العلوم الاجتماعية بأنه مصدر معقول لوصف مدى قوة العلاقة المصاحبة لمقادير «ر» المبينة.

---

(١) Levin and Fox, (1988, p. 311).



الشكل رقم (٦,١). علاقة طردية قوية.

الشكل رقم (٦,٢). علاقة سلبية قوية.

الشكل رقم (٦,٣). علاقة طردية خفيفة.

الشكل رقم (٦,٤). علاقة سلبية خفيفة.

الشكل رقم (٦,٥). علاقة غير خطية.

الشكل رقم (٦,٦). لا توجد أية علاقة.

الجدول رقم (٦,١). وصف مدى قوة العلاقة المصاحبة لمقادير « ر ».

مقدار «ر»	وصف قوة واتجاه العلاقة
-١,٠٠	ارتباط تام عكسي (سالب)
-٠,٦٠	ارتباط قوي عكسي
-٠,٣٠	ارتباط متواضع عكسي
-٠,١٠	ارتباط ضعيف عكسي
٠,٠٠	لا يوجد ارتباط
+٠,١٠	ارتباط ضعيف طردي
+٠,٣٠	ارتباط متواضع طردي
+٠,٦٠	ارتباط قوي طردي
+١,٠٠	ارتباط تام طردي

٦,٢ حساب معامل ارتباط بيرسون «ر»

يشترط لاستخدام معامل ارتباط بيرسون لقياس درجة العلاقة بين متغيرين (س ، ص) – كما أسلفنا القول – أن يقاس كل منهما على مستوى القياس الفاصل (أو الفتري)، وأن تكون العلاقة بينهما خطية، وأن يكون توزيعهما طبيعيا، وأن يكون اختيار مفردات الدراسة اختيارا عشوائيا (أي أن تكون أزواج القيم س، ص مستقلة بعضها عن بعض). ولقد أخذ العلم بأن معامل ارتباط بيرسون يعتمد في تحديد مقداره واتجاه مساره على مقارنة أزواج القيم؛ فإذا كانت العلاقة بين المتغيرين طردية فإن القيم الكبيرة للمتغير (س) تقترن بالقيم الكبيرة للمتغير (ص) والقيم الصغيرة للمتغير (س) تقترن بالقيم الصغيرة للمتغير (ص). أما إذا كانت العلاقة عكسية (أو سلبية) بين المتغيرين، فإن القيم الكبيرة للمتغير (س) تقترن بالقيم الصغيرة للمتغير (ص) والعكس بالعكس. أما قوة أو ضعف العلاقة بين المتغيرين فتتحددان بمدى قرب أو بعد قيمة «ر» من الواحد الصحيح مثلما سبق شرحه في القسم ٦,١ .

ويمكن الحصول على قيمة معامل ارتباط بيرسون «ر» من تعويض القيم في واحدة من المعادلتين الأشهر التاليتين:

(١) $$\text{ر} = \frac{\text{مجـ}(\text{س} - \bar{\text{س}})(\text{ص} - \bar{\text{ص}})}{\sqrt{\text{مجـ}(\text{س} - \bar{\text{س}})^{٢} \times \text{مجـ}(\text{ص} - \bar{\text{ص}})^{٢}}}$$ (٣٠)

$$(٢) \quad ر = \frac{ن \, مجـ(س \, ص) - (مجـ س)(مجـ ص)}{\sqrt{\left[ن \, مجـ س^{٢} - (مجـ س)^{٢}\right]\left[ن \, مجـ ص^{٢} - (مجـ ص)^{٢}\right]}} \quad (٣١)$$

حيث ن = عدد المفردات تحت الدراسة، سَ، صَ المتوسط الحسابي لـ (س) و (ص) على التوالي.

مثال (٦,١)

لنأخذ الآن مثالا تطبيقيا على إيجاد معامل الارتباط «ر». هب أن الجدول رقم (٦,٢) يعكس بيانات مجموعة مكونة من خمسة «٥» موظفين عن الدخل الشهري وعدد سنوات الدراسة لكل منهم. أوجد قيمة «ر» باستخدام القاعدة (٣٠) ثم باستخدام القاعدة (٣١) وعلق على النتيجة.

الجدول رقم (٦,٢). توزيع خمسة موظفين حسب عدد سنوات التعليم والدخل الشهري.

عدد سنوات التعليم	٦	٩	١٢	١٦	٢٠
الدخل الشهري (٠٠٠) ريالا	١	٣	٤	٦	١٠

ولحل المسألة باستخدام القاعدة (٣٠) نرسم الجدول رقم (٦,٣) الذي يوضح الأعمدة الضرورية التي تساعد على الحل.

الجدول رقم (٦,٣). إيجاد معامل الارتباط «ر» باستخدام القاعدة (٣٠).

س	ص	(س − سَ) = (س − ١٢,٦)	(ص − صَ) = (ص − ٤,٨)	(س − سَ)(ص − صَ)	(س − سَ)^٢	(ص − صَ)^٢
٦	١	-٦,٦	-٣,٨	٢٥,٠٨	٤٣,٥٦	١٤,٤٤
٩	٣	-٣,٦	-١,٨	٦,٤٨	١٢,٩٦	٣,٢٤
١٢	٤	-٠,٦	-٠,٨	٠,٤٨	٠,٣٦	٠,٦٤
١٦	٦	+٣,٤	+١,٢	٤,٠٨	١١,٥٦	١,٤٤
٢٠	١٠	+٧,٤	+٥,٢	٣٨,٤٨	٥٤,٧٦	٢٧,٠٤
٦٣	٢٤			٧٤,٦٠	١٢٣,٢	٤٦,٨

وبالتعويض في القاعدة (٣٠) :

$$ر = \frac{٧٤,٦٠}{\sqrt{٤٦,٨ \times ١٢٣,٢}}$$

$$= \frac{٧٤,٦٠}{\sqrt{٥٧٦٥,٧٦}}$$

$$= \frac{٧٤,٦٠}{٧٥,٩٣٢٦}$$ = ٠,٩٨ ← علاقة طردية قوية جدا بين المتغيرين.

والآن نحل نفس المثال (١ , ٦) بتطبيق القاعدة (٣١) فنرسم الجدول رقم (٤ , ٦) المساعد على الحل.

الجدول رقم (٦.٤). إيجاد معامل الارتباط «ر» باستخدام القاعدة (٣١).

س	ص	س ص	$س^2$	$ص^2$
٦	١	٦	٣٦	١
٩	٣	٢٧	٨١	٩
١٢	٤	٤٨	١٤٤	١٦
١٦	٦	٩٦	٢٥٦	٣٦
٢٠	١٠	٢٠٠	٤٠٠	١٠٠
٦٣	٢٤	٣٧٧	٩١٧	١٦٢

والآن نطبق القاعدة (٣١) وبالتعويض فيها نحصل على :

$$ر = \frac{٥ \times ٣٧٧ - ٦٣ \times ٢٤}{\sqrt{[٤٥٨٥ - ٣٩٦٩][٨١٠ - ٥٧٦]}}$$

$$= \frac{٣٧٣}{\sqrt{٦١٦ \times ٢٣٤}}$$

$$= \frac{٣٧٣}{\sqrt{١٤٤١٤٤}}$$

$$= \frac{٣٧٣}{٣٧٩,٦٦} = ٠,٩٨ \leftarrow$$ علاقة طردية قوية جدا بين المتغيرين . وبمعنى آخر ، فإنه كلما زادت عدد سنوات التعليم زاد الدخل الشهري للفرد في هذه المجموعة .

٦,٣ حساب معامل ارتباط سبيرمان للرتب «$ر_ن$»

إذا كان الباحث في حاجة إلى قياس قوة واتجاه العلاقة بين متغيرين أحدهما أو كلاهما مقاس على ميزان القياس الرتبي يمكنه استخدام معامل ارتباط سبيرمان للرتب «$ر_ن$» الذي سبقت الإشارة إليه . والمقصود بالرتب هنا هو إيجاد رُتب (أو ترتيب) القراءات (س ، ص) مع بقاء كل قراءة مكانها وذلك بإجراء ترتيب القراءات إما تصاعديا أو تنازليا ولقد تدربنا على ذلك في الفصول السابقة ولكن لا بأس من إعطاء مثال . إذا طلب منّا مثلا وضع الرتب اللازمة للأرقام : ٧ ، ١٨ ، ٨ ، ١٠ ، ١٥ فما علينا إلا أن نجري ترتيبا تصاعديا (أو تنازليا بالطبع) لها كما يلي :

① ② ③ ④ ⑤
٧ ، ٨ ، ١٠ ، ١٥ ، ١٨

ثم نعطي أصغر رتبة لأصغر قيمة في الترتيب ، والرتبة التي تليها للقيمة التي تلي القيمة الصغرى . . . وهكذا ، فنتحصل على الجدول رقم (٦,٥) .

الجدول رقم (٦,٥). إيجاد الرتب اللازمة للقراءات «س» السابقة.

س	٧	١٨	٨	١٠	١٥
رتب س	١	٥	٢	٣	٤

إذا تم ترتيب قيم مجموعة من المفردات ترتيبا تصاعديا أو تنازليا (١ ، ٢ ، ٣ ، . . . ، ن) بالنسبة لكل متغير من المتغيرين س، ص فإنه يمكن دراسة الارتباط بينهما باستخدام القاعدة:

$$ر_{ت} = ١ - \frac{٦ \text{ مجـ } ف^{٢}}{ن\left(ن^{٢} - ١\right)} \qquad (٣٢)$$

حيث:

$ر_{ت}$ = معامل ارتباط سبيرمان للرتب.

ف = الفرق بين رتبتي المفردة نفسها في المتغيرين س، ص.

ن = عدد المفردات (أو عدد أزواج الرتب).

ويجب ملاحظة أن قيمة مجـ ف = صفرا على الدوام، إلا أن قيمة مجـ $ف^{٢}$ = عددا موجبا (+).

مثال (٦,٢)

أوجد معامل ارتباط بيرسون «$ر_{ت}$» لدرجات خمسة طلاب كما يوضحها الجدول رقم (٦ , ٦)، حيث تمثل س درجاتهم في الجزء النظري من المادة وتمثل ص درجاتهم في الجزء العملي منها.

الجدول رقم (٦,٦). درجات خمسة طلاب في الجزء النظري للمادة (س) والجزء العملي (ص).

س	٨	١٢	١٥	١٤	١٢
ص	١٠	١١	١٤	١٤	١١

وكما اعتدنا في حل المسائل المشابهة في الأمثلة السابقة، نقوم برسم الجدول رقم (٦ , ٧) الذي يوضح الأعمدة الضرورية لحساب الخانات التي سوف نستخدمها في تفعيل القاعدة (٣٢).

الجدول رقم (٦،٧). إيجاد معامل ارتباط بيرسون «$ر_ت$» للخمسة طلاب.

س	ص	رتبة س	رتبة ص	ف	ف٢
٨	١٠	١	١	صفر	صفر
١٢	١١	٢,٥	٢,٥	صفر	صفر
١٥	١٤	٥	٤,٥	٠,٥	٠,٢٥
١٤	١٤	٤	٤,٥	-٠,٥	٠,٢٥
١٢	١١	٢,٥	٢,٥	صفر	صفر
					٠,٥٠

والآن بالتعويض في القاعدة (٣٢) مع ملاحظة أن «ن» = ٥ :

$$ر_ت = ١ - \frac{٦ \times ٠,٥}{٥ (٢٥ - ١)}$$

$$= ١ - \frac{٣}{١٢٠}$$

= ١ – ٠,٠٢٥ = ٠,٩٨ تقريبا .

أي أن هناك ارتباطا قويا جدا وفي الاتجاه الموجب بين درجات الطلاب في الجزء النظري ودرجاتهم في الجزء العملي .

مثال (٦،٣)

أوجد معامل ارتباط بيرسون «$ر_ت$» بين الحالة الاجتماعية لأسرة الزوج وأسرة الزوجة في عينة من سبع أسر بياناتهم كما هي موضحة بالجدول رقم (٦,٨) .

الجدول رقم (٦،٨). الحالة الاجتماعية لعينة من سبع أسر.

حالة أسرة الزوج	فقيرة	متوسطة	غنية	غنية جدا	متوسطة	فقيرة جدا	متوسطة
حالة أسرة الزوجة	متوسطة	غنية	متوسطة	غنية جدا	متوسطة	فقيرة	غنية جدا

نكون الجدول رقم (٩ , ٦) لتسهيل مهمة الحل كما هو معتاد.

الجدول رقم (٩ , ٦). إيجاد معامل ارتباط بيرسون «$ر_{ت}$» بين السبع أسر.

حالة أسرة الزوج	حالة أسرة الزوجة	رتبة أسرة الزوج	رتبة أسرة الزوجة	ف	ف^{٢}
فقيرة	متوسطة	٢	٣	١-	١
متوسطة	غنية	٤	٥	١-	١
غنية	متوسطة	٦	٣	٣	٩
غنية جدا	غنية جدا	٧	٦,٥	٠,٥	٠,٢٥
متوسطة	متوسطة	٤	٣	١	١
فقيرة جدا	فقيرة	١	١	صفر	صفر
متوسطة	غنية جدا	٤	٦,٥	٢,٥-	٦,٢٥
–	–	–	–	–	١٨,٥

والآن نعوض موجودات الأعداد المبينة بالجدول رقم (٩ , ٦) في المعادلة (٣٢) لنتحصل على:

$$ر_{ت} = ١ - \frac{٦ \times ١٨,٥}{٧(٤٩ - ١)}$$

$$= ١ - \frac{٦ \times ١٨,٥}{٧ \times ٤٨}$$

$$= ١ - \frac{١١١}{٣٣٦}$$

= ١ - ٠,٣٣ = ٠,٦٧ أي أن هناك ارتباطا قويا بين حال أسرة الزوج وحال أسرة الزوجة في الاتجاه الموجب.

٦,٤ أسئلـة

١- البيانات التالية توضح درجات عشرة طلاب في اختبار مادة التاريخ «س» ودرجاتهم في اختبار مادة الجغرافيا «ص» .

س	٩٠	٩٥	٧٠	٧٠	٦٥	٦٥	٦٥	٤٠	٥٥	٦٠
ص	٩٧	٩٧	٨٥	٦٥	٧٠	٧٠	٦٠	٥٥	٤٠	٧٠

(أ) أوجد معامل ارتباط بيرسون «ر» بين المتغيرين س، ص .

(ب) أوجد معامل ارتباط سبيرمان للرتب «رن» بين س، ص .

(جـ) علق على النتيجة التي تحصلت عليها في كل من (أ) و(ب) .

(د) أيهما أنسب استخداما لقياس العلاقة بين هذين المتغيرين : «ر» أم «رن» ولماذا؟

٢- فيما يلي طول الأم بالسنتيمتر س، وطول ابنتها بالسنتيمتر، ص .

س	١٦٧,٥	١٦٠	١٥٥	١٦٢,٥	١٧٢,٥	١٥٧,٥	١٦٢,٥	١٦٥
ص	١٧٥	١٧٢,٥	١٦٢,٥	١٧٠	١٦٥	١٥٠	١٦٠	١٦٥

(أ) ارسم شكل الانتشار لبيانات الجدول أعلاه .

(ب) احسب معامل ارتباط بيرسون باستخدام معادلتين تعرفهما .

(جـ) احسب معامل ارتباط سبيرمان للرتب .

٣- عدد شروط استخدام معامل ارتباط بيرسون فيما يتعلق بطبيعة البيانات التي تتم معالجتها وطبيعة المجتمع الذي تمثله .

٦,٥ اصطلاحات ينبغي تذكرها

- الارتباط .
- قوة العلاقة .
- اتجاه العلاقة (سلبي ، إيجابي) .
- معامل الارتباط .
- شكل الانتشار (أو مخطط الانتشار) .

الفصل السابع

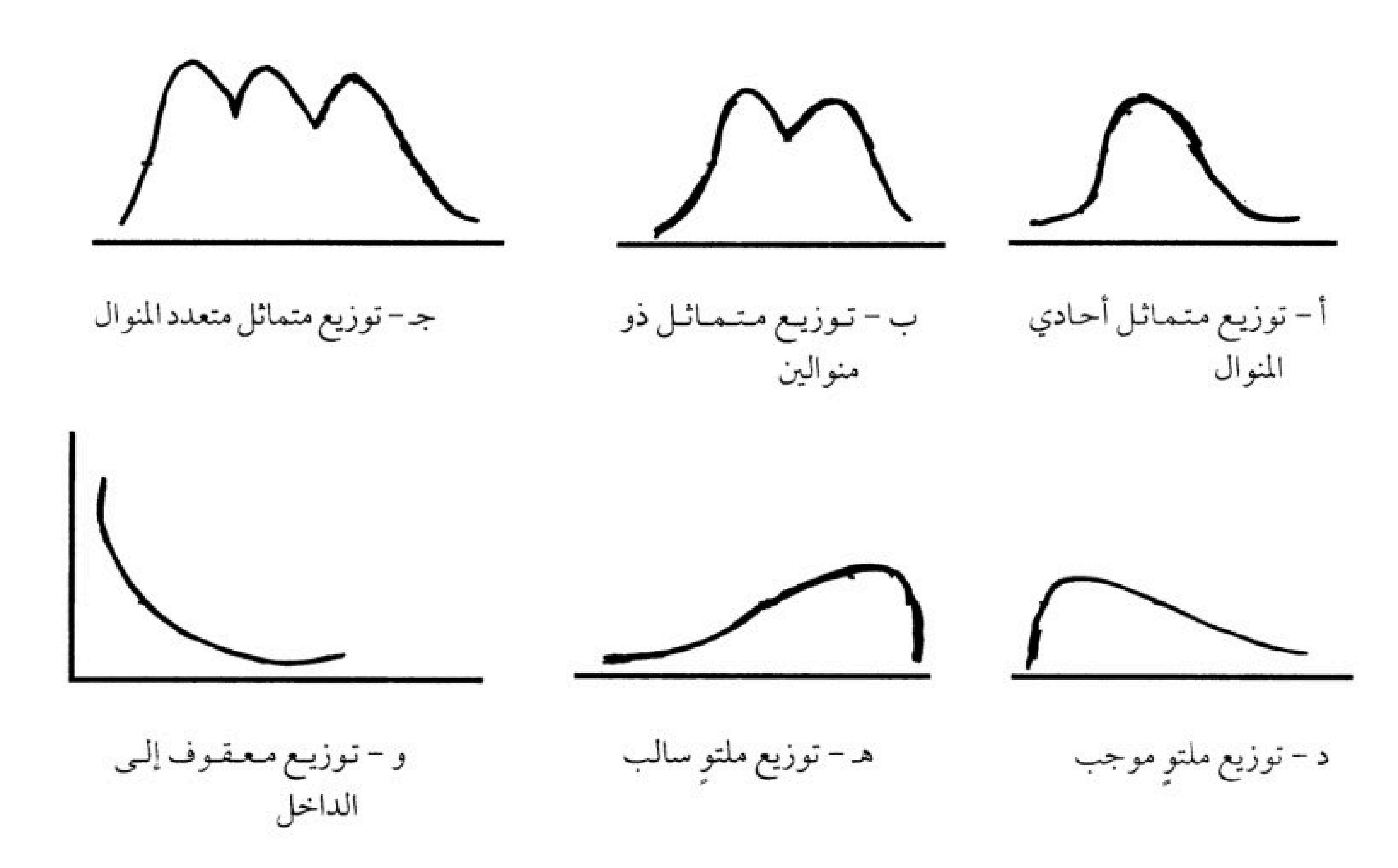
دور الاختبار الإحصائي «مربع كاي» (Chi-Square) في اختبارات فروض البحوث الاجتماعية

٧,١ تقديـــم

لقد تعلمنا في الفصل الثالث (عرض البيانات) أن التوزيع التكراري للقيم يمكن تمثيله بيانيا باستخدام المضلع التكراري الذي يمكن بالتالي تمهيد الخط الذي يصل نقاطه المختلفة يدويا لنحصل على المنحنى التكراري. والشكل العام للمنحنى التكراري للمجتمع يعكس التوزيع النظري لقيم المتغير المتصل. ولقد استفاد الإحصائيون - بعد معايرة standardization المساحة التي تنحصر بين أي منحنى تكراري والمحور الأفقي وجعلها تمثل الوحدة (أي الواحد الصحيح) - من التوزيع النظري لتلك القيم في اختبارات الفروض المختلفة tests of different hypotheses وكذلك في إمكانية إجراء المقارنات بين التكرارات الملاحظة[١] والتكرارات المتوقعة تبعا للتوزيع النظري في المجتمع الذي سحبت منه مفردات الدراسة بُغية دراسة مدى تمثيل العينة للمجتمع الذي تم سحبها منه. ويلاحظ الدارس أن بالإمكان تمثيل تلك التوزيعات التكرارية الملاحظة بمنحنيات تكرارية متنوعة وبحسب طبيعة البيانات الخاصة بالعينة. فبينما نجد أن هناك منحنيات لتوزيعات تكرارية مركزية أو متماثلة (أي أن المنحنى يمكن تطابق نصفيه) نجد أخرى لتوزيعات تكرارية غير متماثلة حيث يكون المنحنى الخاص بمثل

(١) التكرارات الملاحظة نعني بها التوزيع التكراري للعينة، وسميت التكرارات الملاحظة لأن درجاتها لوحظت علميا (وهي درجات مفردات العينة).

هذه التوزيعات مائلا إلى إحدى الجهتين (أي لا يتطابق نصفاه). وتختلف هذه التوزيعات المتماثلة وغير المتماثلة باختلاف ارتفاع أو انخفاض المنحنى عن القاعدة (أي المحور الأفقي) وبحسب عدد المنوالات modes، وطول كل من جهتي المنحنى. والأشكال أرقام (١ ,٧ أ - و) تمثل بعض النماذج لمنحنيات التوزيعات التكرارية.



الشكل رقم (٧,١). بعض النماذج لمنحنيات التوزيعات التكرارية.

وبالطبع، فإن هناك منحنيات لتوزيعات تكرارية خاصة بالمجتمع ويعتبر التعرف عليها مهما جدا في الإحصاء الاستدلالي inferential statistics حيث إن التعرف على خصائص المجتمع في ضوء دراسة خصائص العينة المسحوبة منه تتطلب معرفة شكل توزيع ذلك المجتمع. ولصعوبة، بل وربما لاستحالة حصر جميع مفردات المجتمعات اكتفى العلماء المختصون بمعالجة هذه المشكلة استنادا إلى التوصل إلى معادلات رياضية خاصة بمختلف المجتمعات بغية تحديد أشكال تلك التوزيعات نظريا، ولذلك أطلق على توزيعات المجتمعات المستنبطة بهذه الكيفية اسم «التوزيعات النظرية» أو «التوزيعات المتوقعة»، ومن أشهرها التوزيع الطبيعي normal distribution والذي

يشبه الشكل رقم (١ , ٧أ) . أما في الحالات[٢] التي لا يستطيع فيها الباحث الجزم بشكل توزيع المجتمع الذي يدرس مفردات تنتمي إليه ، أو عندما تكون هذه المفردات قليلة[٣] بالنسبة إلى حجم المجتمع الذي سحبت منه ، أو أن العينة المسحوبة لم تكن عينة احتمالية ، أو أن البيانات التي يتعامل معها بيانات نوعية وليست كمية تقتضي التعامل مع درجات المفردات ، فإن من أشهر التوزيعات النظرية الإحصائية التي يعتمد عليها في اختبارات الفروض هو توزيع مربع كاي [chi-Square (coded χ^2)] ويرمز إليه بالرمز «كا٢» . وهو مهم بصفة خاصة في البحوث الاجتماعية التطبيقية التي تكثر فيها الحالات سابقة الذكر والتي يكون فيها استخدام مربع كاي مناسبا . ويشبه المنحنى الذي يوضح توزيع مربع كاي «بدرجة حرية واحدة»[٤] الشكل رقم (١ , ٧و) إلا أنه يقترب من شكل التوزيع الطبيعي كلما زاد عدد مفردات المجتمع (نظريا) عن مفردتين اثنتين .

في كثير من البحوث الاجتماعية نجد أن القضية المدروسة تتطلب من المبحوث إبداء رأيه حولها باختيار بديل واحد أو أكثر من بديل واحد من بين مجموعة من البدائل المدونة سلفا كإجابات لسؤال يتطلب الإجابة «بنعم» أو «لا» أو «لا أدري» مثلا ، أو «أوافق بشدة» ، «أوافق» ، أوافق إلى حد ما» ، «لا رأي لي» ، «لا أوافق بشدة» ، «لا

---

(٢) في جميع هذه الحالات يتم الاعتماد على اختبارات البيانات اللامعلمية والتي من أهمها اختبار مربع كاي (كا٢) . ومن هذه الاختبارات أيضا اختبار «ذو الحدين» واختبار «كولمجروف - سميرنوف» لعينة واحدة ، اختبارات «ماكنمار» ، «ولكوكسن» و«الإشارة» للمقارنة بين عينتين مترابطتين ، اختبار «فشر» ، اختبار «مان-وتني» ، اختبار «الوسيط» واختبار «كولموجروف-سميرنوف» للمقارنة بين عينتين مستقلتين ، اختبار «كوجران» واختبار «فريدمان-تحليل التباين من الدرجة الثانية» للمقارنة بين عدة عينات مترابطة ، اختبار «كروسكال-واليز» للمقارنة بين عدة عينات مستقلة وكذلك اختبار «الوسيط» ، إضافة إلى بعض معاملات الارتباط للبيانات الاسمية والرتبية مثل معامل ارتباط «فاي» ، «سبيرمان للرتب» و«كاندل تاو للرتب» . أما أشهر اختبارات البيانات المعلمية فهي اختبار «ت» واختبار «ف» ومختلف أنواع تحليل التباين . ويجدر بالذكر أن كلمة «معلمية» parametric تشير إلى أن المقياس الإحصائي (كالمتوسط والانحراف المعياري) يخص المجتمع فيما تشير كلمة «لا معلمية» non-parametric إلى أن المقياس الإحصائي يخص العينة sample .

(٣) أقل من «٥٠» مثلا .

(٤) «درجة الحرية» اصطلاح إحصائي وهو يساوي عدد المفردات المدروسة «ن» (مع معالجة خاصة عند جدولة البيانات في جدول مزدوج) في حالة توزيع مربع كاي .

أوافق»، «لا أوافق إلى حد ما». . وهلم جرا. وقد يتعلق أمر البحث هنا بمعرفة مدى الفروق الموجودة بين هذه التكرارات الملاحظة والتكرارات المتوقع الحصول عليها بشأن الاستجابات نفسها في المجتمع الأصلي. ولمعرفة هذه الفروق وما إذا كانت فروقا جوهرية أم ناتجة عن الصدفة تجري مقارنة التكرارات الملاحظة مع التكرارات المتوقعة (وهي التكرارات النظرية التي تم التعرف عليها نتيجة توقعات أخرى غير مبنية على الدراسة الميدانية) مما يستدعي استخدام التوزيع النظري لمربع كاي. وإذا أضفنا إلى هذا، الحالات التي يجد الباحث فيها نفسه غير قادر على استخدام التوزيع الطبيعي لاختبار فروض دراسته[٥]، فإن توزيع مربع كاي يكون هو البديل الأنسب لحل مشكلة الباحث الاجتماعي. وسوف نتعرض، بعد عرض موجز لمفهوم اختبارات الدلالة الإحصائية، إلى كيفية التطبيق العملي لـ(كا٢) آخذين في الاعتبار أسس تطبيقه في مختلف العينات المدروسة التي تختلف باختلاف درجة ترابطها.

٧,٢ مفهوم اختبار الدلالة الإحصائية

لقد تطرقنا بإسهاب إلى معنى الإحصاء الاستدلالي وصياغة فرض البحث في أبسط صوره وذلك في الفصل الأول من فصول هذا المؤلف. ولقد ذكرنا أن الإحصاء الاستدلالي يُعنى بطرق تعميم نتائج العينة على المجتمع الذي تم سحبها منه. والاهتمام بالتعميم ناتج من أن الباحث الاجتماعي يهدف في النهاية إلى كشف الحقائق عن المجتمعات ذاتها وليس عن جزئية صغيرة منها كالعينة، لأن التخطيط الاجتماعي لصالح المجتمع ورفاهيته يسعى لأن يشمل الكل وليس الجزء. وبإيجاز شديد، فإن ما يسمى «باختبار الدلالة الإحصائية» test of statistical significance هو أسلوب إحصائي يسعى لإثبات ما إذا كانت النتائج المتحصل عليها من دراسة العينة تنطبق على المجتمع الأم لهذه العينة بدرجة ثقة (إحصائية) تفوق اعتبارات الصدف أم لا. ومن متطلبات هذا الإثبات أن يكون هناك ما يسمى «بفرض العدم» null hypothesis (أو الفرض الصفري

---

(٥) إذا سحبت عينة احتمالية من مجتمع ذي توزيع طبيعي، والبيانات التي تخص هذه العينة يمكن قياسها على ميزان القياس الفاصل (أو الفتري)، فإن الاختبارات الإحصائية المعتادة لدراسة الفروض في مثل هذه الحالات تشمل اختبار «ت» واختبار «ف» وطرق تحليل التباين وما إليها.

في بعض المؤلفات) وهو نقيض فرض البحث research hypothesis (أو الفرض العلمي في بعض المؤلفات) الذي سقنا أمثلة له في الفصل الأول، ومنها: كلما زادت مدة بقاء الوالدين مع أبنائهما، قلت مخاطر انحراف هؤلاء الأبناء. وفرض العدم الذي ينطلق منه اختبار الدلالة الإحصائية لفرض البحث في هذا المثال يصاغ هكذا: «لا تقل مخاطر انحراف الأبناء بازدياد مدة بقاء والديهم معهم». وعلى اختبار الدلالة الإحصائية أن يثبت عكس هذه المقولة ليؤكد صحة فرض البحث، وهذا هو الشئ الذي يسعى إليه الباحث. وإذا فشل اختبار الدلالة الإحصائية في رفض فرض العدم (أو الفرض الصفري) فستكون المحصلة أن النتيجة المتحصل عليها من دراسة العينة لا تنطبق على المجتمع الأم بدرجة ثقة محددة (متفق عليها سلفا) وإن كانت تلك النتيجة تطابق رؤى الباحث حيث سيكون ذلك ناتج عن محكات الصدفة فقط حسب الأجندة العلمية لاختبار الدلالة الإحصائية المستخدم وفقا لمستوى الدلالة (أو مستوى المعنوية في بعض المؤلفات) المختار.

إن مستوى الدلالة المشار إليه في الفقرة السابقة لا يعدو كونه يعبر عن هامش الخطأ المسموح به في تعميم نتائج العينة على المجتمع المسحوبة منه. ويحدد هامش الخطأ هذا في معظم الأبحاث الاجتماعية بـ٥٪ وأحيانا ١٪. وتفسير هذه النسبة يتمثل في أننا نقبل أن نكون مخطئين في تعميمنا لنتائج العينة على المجتمع (أي مخطئين في رفضنا فرض العدم) خمس مرات من كل مئة مرة نسحب فيها عينة من مجتمع ما في حالة تحديد ٥٪ لهامش الخطأ، أو مرة واحدة في كل مئة مرة في حالة تحديد هامش الخطأ (أو مستوى الدلالة) بـ١٪. ومن المعروف أن درجة الثقة المشار إليها في الفقرة السابقة لا تعدو كونها المكمل لمستوى الدلالة إلى مئة. أي أنه في حالة أن مستوى الدلالة الذي تم اختياره لاختبار الدلالة الإحصائية لنتائج عيناتنا هو ٥٪ (أو «٠٥ , ٠») فإن درجة الثقة تكون ١٠٠ - ٥ = ٩٥٪. وبمعنى آخر فإننا نثق بمعدل ٩٥ مرة من كل مئة مرة نسحب فيها عينة ما من مجتمع ما، في أن نتائج تلك العينة تماثل النتائج التي سوف يتحصل عليها من المجتمع المعني ككل من الناحية النظرية.

٧.٣ استخدام (كا٢) في اختبار الفروق لعينتين مستقلتين أو أكثر

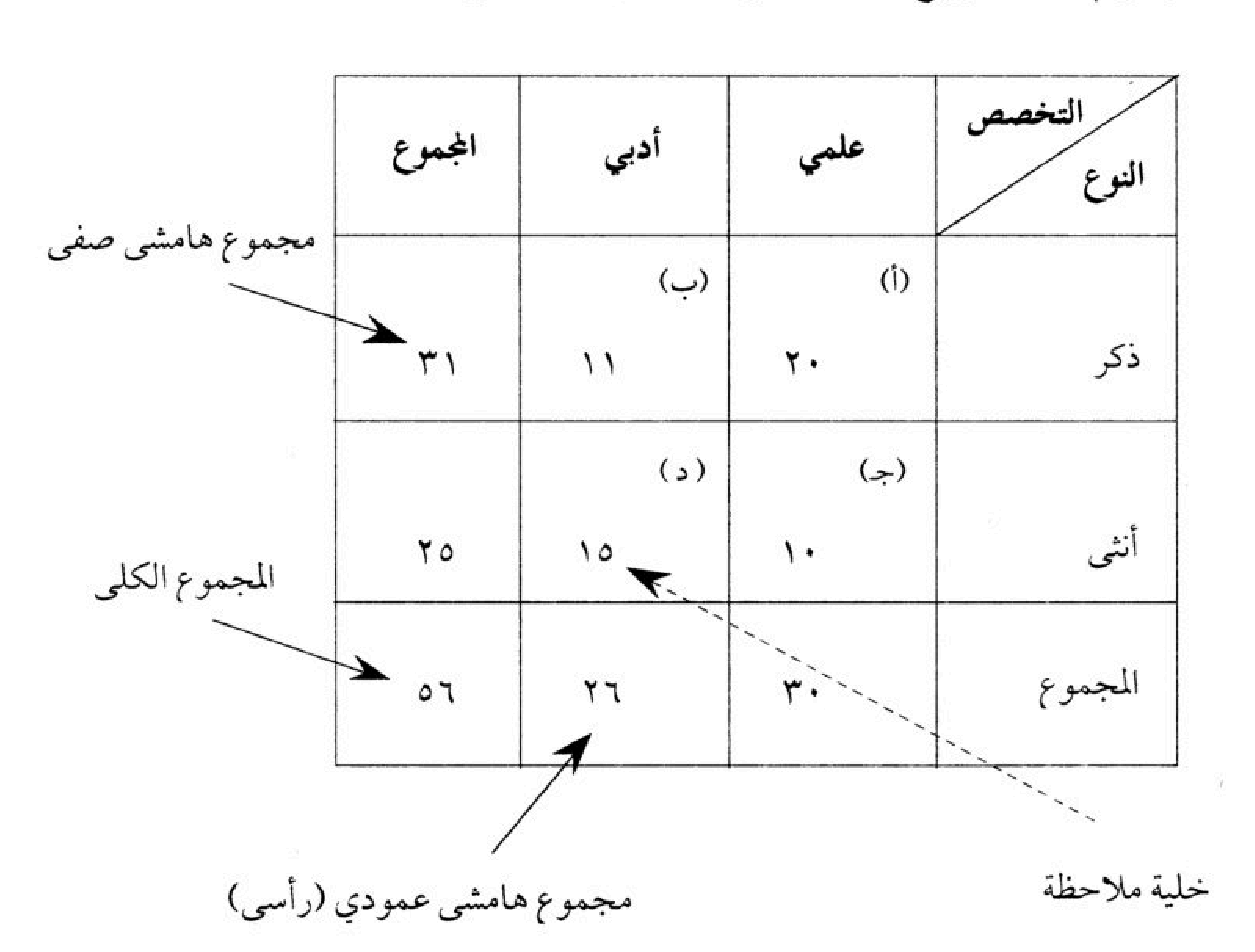
هب أن باحثا أراد أن يتأكد عما إذا كانت هناك «فروق جوهرية» بين خيارات الطلاب الذكور وخيارات الطالبات فيما يتعلق بنوع التخصص (علمي أو أدبي) الذي

يرغبون الالتحاق به عند دخولهم المدارس الثانوية وهم بعد في الصف النهائي بالمرحلة المتوسطة . وعبارة فروق جوهرية تم تقويسها للتنبيه إلى أنها كثيرا ما ترد في التعليق على نتائج اختبارات الدلالة الإحصائية وهي تعني «فروق ذات دلالة إحصائية» . والعبارتان تستخدمان في تبادل رتيب وهما تعنيان الشيء ذاته ولا فرق إطلاقا بينهما في المعنى . ولنأخذ المثال (١ , ٧) لمعرفة كيفية حساب هذه الفروق ، وإيجاد قميتي كا[٢] المحسوبة الجدولية ومقارنتهما لقبول فرض أو رفضه .

مثال (٧،١)

هب أن هذا الباحث اختار عينتين عشوائيتين إحداهما من الطلاب والأخرى من الطالبات وسجل استجابات أفراد العينتين فيما يتعلق برغبة الالتحاق بنوع معين من نوعي التخصص فنظّم تلك الاستجابات في الجدول رقم (١ , ٧) . ولدى الباحث

الجدول رقم (٧،١). توزيع لـ ٥٦ طالبا وطالبة حسب التخصص.



التخصص / النوع	علمي	أدبي	المجموع
ذكر	(أ) ٢٠	(ب) ١١	٣١
أنثى	(جـ) ١٠	(د) ١٥	٢٥
المجموع	٣٠	٢٦	٥٦

الآن ما يلي من إقرارات :

فرض العدم : ليست هناك فروق ذات دلالة إحصائية بين استجابات الطلاب الذكور والطالبات تجاه التخصص الذي يرغبونه .

فرض البحث : إن استجابات الذكور تجاه التخصص الذي يرغبونه تختلف عن استجابات الطالبات نحو التخصص الذي يرغبنه .

مستوى الدلالة : ٠,٠٥

إن ما يهم الباحث الآن وقد جدول بياناته في جدول توافق «٢ × ٢» (أي عمودين في صفين ، مع ملاحظة أن العمود والصف الخاصين بالمجاميع الهامشية لا يحسبان عند تعريف الجدول بعدد الأعمدة وعدد الصفوف) وحدد فرض البحث وتلقائيا حدد أيضا فرض العدم كما حدد مستوى الدلالة الذي سوف يرفض أو يقبل عنده فرض العدم ، هو أن يحسب قيمة كا^{٢} من التكرارات الملاحظة أمامه في جدوله التكراري الذي أنشأه من بياناته الخام التي جمعها . ثم يقارن قيمة كا^{٢} التي حسبها من بياناته مع قيمة كا^{٢} النظرية التي يتم تحديدها بتعيين القيمة التي يتقاطع عندها مستوى الدلالة الذي حدده بـ ٠,٠٥ مع عدد درجات الحرية المقترنة بجدوله الذي أنشأه ، وذلك بالبحث عن تلك القيمة في جدول التوزيع النظري لـ (كا^{٢}) والذي أنجزه الإحصائيون من قبل ، وتوجد نسخ منه في معظم الكتب المنهجية في الإحصاء الاستدلالي ويطلق عليها في بعض مؤلفات الإحصاء «كا^{٢}» الجدولية إشارة إلى جدول التوزيع النظري لـ (كا^{٢}) الذي غالبا ما يوجد مثبتا كملحق في مؤلفات الإحصاء الاستدلالي كما ذكرنا ذلك سلفا . وقبل أن نتحدث عن الهدف من مقارنة قيمتي كا^{٢} المحسوبة والجدولية ، علينا أن نحدد كيف نجد عدد درجات الحرية للجدول رقم (٧,١) والتي نحتاجها لتتضافر مع مستوى الدلالة المختار (وهو هنا ٠,٠٥) فتحددان لنا قيمة كا^{٢} الجدولية كما سبقت الإشارة إلى هذا الأمر . ولإيجاد درجات الحرية من جدول كهذا أو أي جدول آخر يعرف بالرمز المزدوج (و × هـ) حيث تشير «و» إلى عدد الصفوف وتشير «هـ» إلى عدد الأعمدة ؛ تطبق القاعدة :

عدد درجات الحرية = (و - ١) (هـ - ١) (٣٣)

وبتطبيق القاعدة (٣٣) على الجدول رقم (١ , ٧) نحصل على :

عدد درجات الحرية = (٢ - ١) (٢ - ١)

= ١ × ١ = ١

وبعد أن نفرغ من حساب قيمة كا٢ من بياناتنا المعطاة في الجدول رقم (١ , ٧) [انظر كيف يتم ذلك بعد قليل]، ونفرغ من إيجاد قيمة كا٢ الجدولية (أو النظرية) بالرجوع إلى جدول التوزيع النظري لـ (كا٢) [انظر الملحق] فإننا نقوم بمقارنة القيمتين لنحدد أيهما أكبر. فإذا كانت قيمة كا٢ المحسوبة تساوي أو تزيد عن قيمة كا٢ الجدولية المقترنة بفرض العدم فإننا نرفض الفرضية الصفرية (أو فرض العدم) وبالتالي نقبل فرض البحث. أما إذا كانت قيمة كا٢ المحسوبة تقل عن قيمة كا٢ الجدولية فإننا نقبل فرض العدم وبالتالي نرفض فرض البحث.

٧,٣,١ حساب قيمة كا٢ من الجدول التكراري

لإيجاد قيمة كا٢ المحسوبة نطبق القاعدة التالية :

$$\text{كا}^2 \text{ (المحسوبة)} = \text{مجـ} \left[\frac{(\text{ل} - \text{قع})^2}{\text{قع}} \right] \qquad (٣٤)$$

حيث :

ل = التكرار «الملاحظ» في أي خلية من خلايا الجدول التكراري.

قع = التكرار «المتوقع» لأي خلية من خلايا الجدول التكراري.

مجـ = حاصل جمع نواتج العمليات لكل خلايا الجدول.

وإذا علمنا أن التكرار الملاحظ موجود في كل خلية من خلايا الجدول فكيف نتحصل على التكرار المتوقع (قع) لأي خلية من خلايا الجدول حتى نتمكن من تطبيق القاعدة (٣٤) لإيجاد قيمة كا٢ المحسوبة؟ وللإجابة على هذا السؤال يستخدم الجدول رقم (١ , ٧) ويحسب التكرار المتوقع (قع) لأي خلية من خلايا أي جدول تكراري (و × هـ) بتطبيق القاعدة التالية :

$$\text{قع} = \frac{\text{(المجموع الهامشى الصفى لصف الخلية)} \times \text{(المجموع الهامشى العمودي لعمود الخلية)}}{\text{المجموع الكلى لتكرارات الجدول}} \qquad (٣٥)$$

وإذا رمزنا إلى المجموع الهامشي الصفي بالرمز (مجصف) وإلى المجموع الهامشي العمودي بالرمز (مجعد)، وإلى المجموع الكلي للمفردات المدروسة بالحرف «ن» فإن القاعدة (٣٥) يمكن أن تكتب :

$$\text{قع} = \frac{\text{مجصف} \times \text{مجعد}}{\text{ن}}$$

ولنجد الآن قيم (قع) للخلايا (أ)، (ب)، (جـ)، (د) في مثالنا أعلاه.

إذن، الخطوة (١) :

$$\text{قع}_{\text{أ}} = \frac{\text{(مجصف أ) (مجعد أ)}}{\text{ن}} = \frac{٣١ \times ٣٠}{٥٦} = ١٦,٦$$

$$\text{قع}_{\text{ب}} = \frac{\text{(مجصف ب) (مجعد ب)}}{\text{ن}} = \frac{٣١ \times ٢٦}{٥٦} = ١٤,٤$$

$$\text{قع}_{\text{جـ}} = \frac{\text{(مجصف جـ) (مجعد جـ)}}{\text{ن}} = \frac{٢٥ \times ٣٠}{٥٦} = ١٣,٤$$

$$\text{قع}_{\text{د}} = \frac{\text{(مجصف د) (مجعد د)}}{\text{ن}} = \frac{٢٥ \times ٢٦}{٥٦} = ١١,٦$$

الخطوة (٢): اطرح التكرار المتوقع من التكرار الملاحظ لكل خلية.

$$\text{ل}_{\text{أ}} - \text{قع}_{\text{أ}} = (٢٠ - ١٦,٦) = ٣,٤$$

$$\text{ل}_{\text{ب}} - \text{قع}_{\text{ب}} = (١١ - ١٤,٤) = -٣,٤$$

$$\text{ل}_{\text{جـ}} - \text{قع}_{\text{جـ}} = (١٠ - ١٣,٤) = -٣,٤$$

$$\text{ل}_{\text{د}} - \text{قع}_{\text{د}} = (١٥ - ١١,٦) = ٣,٤$$

الخطوة (٣): ربع الفروق المتحصل عليها في الخطوة (٢).

$$(\text{ل}_{\text{أ}} - \text{قع}_{\text{أ}})^{٢} = (٣,٤)^{٢} = ١١,٥٦$$

$(ل_{ب} - قع_{ب})^{٢}$ = $(-٤,٣)^{٢}$ = ١١,٥٦

$(ل_{جـ} - قع_{جـ})^{٢}$ = $(-٤,٣)^{٢}$ = ١١,٥٦

$(ل_{د} - قع_{د})^{٢}$ = $(٤,٣)^{٢}$ = ١١,٥٦

الخطوة (٤): قسم كل ناتج متحصل عليه في الخطوة (٣)، على (قع).

$$\frac{(ل_{أ} - قع_{أ})^{٢}}{قع_{أ}} = \frac{١١,٥٦}{١٦,٦} = ٠,٦٩٦$$

$$\frac{(ل_{ب} - قع_{ب})^{٢}}{قع_{ب}} = \frac{١١,٥٦}{١٤,٤} = ٠,٨٠٣$$

$$\frac{(ل_{جـ} - قع_{جـ})^{٢}}{قع_{جـ}} = \frac{١١,٥٦}{١٣,٤} = ٠,٨٦٣$$

$$\frac{(ل_{د} - قع_{د})^{٢}}{قع_{د}} = \frac{١١,٥٦}{١١,٦} = ٠,٩٩٧$$

الخطوة (٥): نجمع النواتج في الخطوة (٤) لنتحصل على قيمة كا^{٢} (المحسوبة).

كا^{٢} (المحسوبة) $= مجـ \left[\frac{(ل - قع)^{٢}}{قع} \right]$

= ٠,٦٩٦

+ ٠,٨٠٣

+ ٠,٨٦٣

+ ٠,٩٩٧

= ٣,٣٥٩

٧,٣,٢ مقارنة كا٢ (المحسوبة) بـ كا٢ (الجدولية)

الخطوة (٦): أوجد عدد درجات الحرية.

عدد درجات الحرية = ١ (انظر المعادلة ٣٣، ص١٥٠).

ولإيجاد كا٢ الجدولية انظر ملحق التوزيع النظري لمربع كاي وحدد تقاطع عدد ١ درجة حرية مع مستوى الدلالة ٠,٠٥ وستجد أن قيمة كا٢ (الجدولية) التي يعينها هذا التقاطع = ٣,٨٤١.

إذن:

الخطوة (٧): قارن بين قيمتي كا٢ (المحسوبة) وكا٢ (الجدولية) لاتخاذ القرار.

كا٢ (المحسوبة) = ٣,٣٥٩

كا٢ (الجدولية) = ٣,٨٤١ بدرجة حرية «١» ومستوى دلالة ٠,٠٥

الخطوة (٨): اتخاذ القرار.

مادام أن كا٢ (المحسوبة) [= ٣,٣٥٩] أقل من كا٢ (الجدولية) [= ٣,٨٤١] فإننا نقبل فرض العدم ونرفض فرض البحث وبذلك تكون النتيجة أنه ليست هناك فوارق ذات دلالة إحصائية بين خيارات مجتمع الطلاب الذكور والطالبات الإناث فيما يتعلق بالتخصص الذي يرغبون الالتحاق به عند دخولهم الثانوي. وأن ارتفاع نسب الذكور في التخصص العلمي عن نسب الإناث المقابلة ما هي إلا نتاج للصدفة المحضة.

٧,٣,٣ قاعدة حسابية مختصرة لـ كا٢ (المحسوبة) من جدول ٢ × ٢

يمكننا تجنب المراحل الطويلة السابقة لحساب كا٢ (المحسوبة) باستخدام القاعدة التالية عندما يكون الجدول التكراري ٢ × ٢، آخذين في الاعتبار تسميات الخلايا السابقة (أ)، (ب)، (جـ)، (د)؛ ن = المجموع الكلي.

$$\text{كا}^2\ (\text{المحسوبة}) = \frac{\text{ن}\ (\text{أ د} - \text{ب جـ})^2}{(\text{أ} + \text{ب})\ (\text{جـ} + \text{د})\ (\text{أ} + \text{جـ})\ (\text{ب} + \text{د})} \qquad (٣٦)$$

وبتطبيق هذه القاعدة [أي القاعدة (٣٦)] على بيانات الجدول رقم (١ ,٧) نتحصل على:

$$\text{كا}^2\ (\text{المحسوبة}) = \frac{٥٦\,(٢٠ \times ١٥ - ١١ \times ١٠)^2}{(٢٠ + ١١)(١٠ + ١٥)(٢٠ + ١٠)(١١ + ١٥)}$$

$$= \frac{٥٦\,(٣٠٠ - ١١٠)^2}{(٣١)(٢٥)(٣٠)(٢٦)}$$

$$= \frac{٥٦ \times (١٩٠)^2}{٦٠٤٥٠٠}$$

$$= \frac{٥٦ \times ٣٦١٠٠}{٦٠٤٥٠٠}$$

$$= \frac{٢٠٢١٦٠٠}{٦٠٤٥٠٠} = ٣,٣٤$$

. وهي نفس القيمة السابقة.

ويجدر بالذكر أن المعادلة (٣٤) لإيجاد قيمة كا^2 (المحسوبة) تنطبق على أي جدول تكراري (و × هـ) حيث يمكن أن يكون عدد الصفوف أكثر من «٢» وكذلك عدد الأعمدة، أي أن المتغيرين المدروس تكرارهما يمكن أن يكونا بفئات تزيد عن الـ«٢» لكل منهما. ولنأخذ مثالا لجدول (٣ × ٣).

مثال (٢ ,٧)

أراد باحث أن يدرس العلاقة بين «مكان التنشئة» و«المستوى التعليمي» لأشخاص بلغو الستين من العمر، إذ يعتقد الباحث بأن هناك فروقا أساسية بين المستويات التعليمية لثلاث مجموعات تم تصنيفها تبعا لخلفيتها وفق مكان التنشئة: حضري، أو ريفي، أو بدوي. قام الباحث بسحب ثلاث عينات عشوائية من المجتمع المعني: ٢٩ من ذوي المنشأ الحضري، و٢٤ من ذوي المنشأ الريفي، و١٨ من ذوي المنشأ البدوي. وكان تصنيفهم وفقا للمستوى التعليمي بفئاته الثلاث: أولي، متوسط، عال قد تم كما هو

موضح في الجدول رقم (٢ , ٧) .

الجدول رقم (٧,٢) توزيع لـ ٧١ شخصاً حسب مكان التنشئة والمستوى التعليمي.

مكان التنشئة / المستوى التعليمي	حضري	ريفي	بدوي	المجموع
تعليم أولى	① ٦	② ٧	③ ١١	٢٤
تعليم متوسط	④ ٩	⑤ ٨	⑥ ٥	٢٢
تعليم عال	⑦ ١٤	⑧ ٩	⑨ ٢	٢٥
المجموع	٢٩	٢٤	١٨	٧١

ولاختبار ما إذا كانت الفروق الملاحظة بين هذه المجموعات الثلاث فيما يتعلق بمستوياتهم التعليمية طبقا لمكان التنشئة فروقا جوهرية أم أنها فقط ناتجة عن الصدفة نبدأ بصياغة الفرض الصفري وفرض البحث كما يلي :

فرض العدم : إن التكرار النسبي لمستوى التعليم الأولي ومستوى التعليم المتوسط ومستوى التعليم العالي هو التكرار النسبي نفسه لكل من الحضر والريف والبدو .

فرض البحث : إن التكرار النسبي لمستوى التعليم الأولي ومستوى التعليم المتوسط ومستوى التعليم العالي ليس هو التكرار النسبي نفسه لكل من الحضر والريف والبدو .

مستوى الدلالة المختار = ٠,٠٥

درجات الحرية = (و − ١) (هـ − ١)
= (٣ − ١) (٣ − ١)
= ٢ × ٢
= ٤ .

بتطبيق القاعدة (٣٣) على الجدول رقم (٢ , ٧)

والآن نشرع في حساب قيمة كا^{2} (المحسوبة) كالمعتاد حيث نبدأ حساب (قع) لكل خلية باستخدام القاعدة (٣٥) قبل أن نطبق القاعدة (٣٤) لإيجاد قيمة كا^{2} المحسوبة. ونسبة لأن تلك الخطوات الطويلة التي اتبعناها في حساب كا^{2} (المحسوبة) كانت لأسباب تعليمية أساسية في هذا المنحنى، سوف نجمل الآن تلك الخطوات في مساحة أقل بما أننا أعدنا تكرار المألوف. نبدأ أولا بحساب التكرارات المتوقعة (قع) لكل خلية كالتالي مع ملاحظة أن الخلايا في الجدول رقم (٢ , ٧) قد رمز إليها بالأرقام الموضحة في دائرة :

الخطوة (١): إيجاد التكرار المتوقع لكل خلية

$$\text{قع}_{١} = \frac{٢٩ \times ٢٤}{٧١} = ٩,٨$$

$$\text{قع}_{٢} = \frac{٢٤ \times ٢٤}{٧١} = ٨,١$$

$$\text{قع}_{٣} = \frac{١٨ \times ٢٤}{٧١} = ٦,١$$

$$\text{قع}_{٤} = \frac{٢٩ \times ٢٢}{٧١} = ٩$$

$$\text{قع}_{٥} = \frac{٢٤ \times ٢٢}{٧١} = ٧,٤$$

$$\text{قع}_{٦} = \frac{١٨ \times ٢٢}{٧١} = ٥,٦$$

$$قع_٧ = \frac{٢٥ \times ٢٩}{٧١} = ١٠,٢$$

$$قع_٨ = \frac{٢٥ \times ٢٤}{٧١} = ٨,٥$$

$$قع_٩ = \frac{٢٥ \times ١٨}{٧١} = ٦,٣$$

الخطوة (٢): إيجاد أجزاء كا٢ (المحسوبة) الخاصة بكل خلية، ثم جمعها باستخدام القاعدة (٣٤)

$$\frac{(٦ - ٩,٨)^٢}{٩,٨} + \frac{(٧ - ٨,١)^٢}{٨,١} + \frac{(١١ - ٦,١)^٢}{٦,١}$$

$$+ \frac{(٩ - ٩)^٢}{٩} + \frac{(٨ - ٧,٤)^٢}{٧,٤} + \frac{(٥ - ٥,٦)^٢}{٥,٦}$$

$$+ \frac{(١٤ - ١٠,٢)^٢}{١٠,٢} + \frac{(٩ - ٨,٥)^٢}{٨,٥} + \frac{(٢ - ٦,٣)^٢}{٦,٣} = كا^٢ \text{(المحسوبة)}$$

كا٢ (المحسوبة) = ١,٤ + ٠,١٥ + ٣,٩ + ٠,٠٠ + ٠,٠٥ + ٠,٠٦ + ١,٤ + ٠,٠٣ + ٢,٩ .
= ٩,٩٦

كا٢ (الجدولية) = ٩,٤٩ (اقتران ٤ درجات حرية ومستوى دلالة ٠,٠٥)

الخطوة (٣): القرار

ما دام أن كا٢ (المحسوبة) [٩,٩٦] أكبر من كا٢ (الجدولية [٩,٤٩] فإننا نرفض فرض العدم ونقبل بالتالي فرض البحث لنقرر أن التكرارات النسبية لمستويات التعليم الثلاثة تختلف اختلافا ذا دلالة إحصائية عما هو متوقع تحت فرض العدم بين مجتمع

الحضر والريف والبدو. وبمعنى آخر فإن التكرارات الملاحظة في الجدول والتي تشي بأن الحضر يظهرون مستوى أعلى في التعليم، يتبعهم في ذلك الريفيون ثم أخيرا البدو تعكس الوضع الحقيقي في المجتمعات الثلاث وليس وضعا يمكن نسبته إلى الصدفة.

٧,٣,٤ تصحيح الخطأ الناتج عن التكرارات المتوقعة الصغيرة الحجم

رغم أن كا٢ كاختبار إحصائي لا معلمي لا يتطلب استخدامه افتراضات كثيرة عن صفات وخصائص العينات وشكل منحنيات توزيع مجتمعاتها كما يتطلب ذلك الاختبارات المعلمية مما حدا بالباحثين في المسائل الاجتماعية تفضيله على غيره من الاختبارات الإحصائية عند تحليل بياناتهم واختبار فرضياتهم، إلا أن لـ كا٢ عيوبه التي من أهمها احتمال الحصول على قيمة لـ كا٢ (المحسوبة) تميل إلى رفض فرض العدم بصورة غير منطقية نتيجة لحساسيتها المفرطة لقيمة التكرار المتوقع (قع). ويتضح هذا جليا إذا تأملنا قاعدة أي مكوّن لـ كا٢ (المحسوبة): $\frac{(\text{ل} - \text{قع})^{2}}{\text{قع}}$ ، إذ إن قيمة هذا المكون تعتمد على حجم البسط والذي يعتمد بدوره على قيمة (قع). فكلما كانت (قع) صغيرة الحجم كان المكوّن كبيرا وتكون المحصلة النهائية لقيمة كا٢ (المحسوبة) أكبر بكثير مما لو كانت قيمة (قع) مماثلة لمعظم مثيلاتها بخلايا الجدول وليست صغيرة بشكل لافت مقارنة بها. ومشكلة مثل هذه تكون مزعجة بصورة خاصة لدى الجداول التكرارية ذات الحجم ٢ × ٢. فقد لوحظ أن تكرارا متوقعا يقل عن ١٠ (عشرة) في أي خلية من الخلايا الأربع التي يتضمنها جدول ٢ × ٢ يسبب مشكلة كهذه ويجب معالجتها قبل التسليم بالقيمة المصححة لـ كا٢ (المحسوبة) بعد المعالجة. أما في الجداول التكرارية ذات الأحجام التي تزيد عن ٢ × ٢ فيتعين ألا يقل التكرار المتوقع في كل خلية لغالبية الخلايا عن ٥ (خمسة) وإلا استوجب ذلك معالجة خاصة أيضا قبل الاعتماد على قيمة كا٢ (المحسوبة) بعلاتها في رفض أو قبول الفرض الصفري.

وسوف ندعم الآن ما ذهبنا إليه في الفقرة السابقة بمثالين.

مثال (٧,٣)

			المجموع
	① ١٥ (١١,٧٦)	② ٥ *(٨,٣٣)	٢٠
	③ ٦ *(٩,٣٣)	④ ١٠ *(٦,٦٧)	١٦
المجموع	٢١	١٥	٣٦

* ← (قع) أقل من ١٠

مثال (٧,٤)

			المجموع
	① ١٠٠ (٩٠)	② ٢٠ (٣٠)	١٢٠
	③ ٣٩ (٤٥)	④ ٢١ (١٥)	٦٠
	⑤ ٩ (٩)	⑥ ٣ #(٣)	١٢
	⑦ ٢ (٦)	⑧ ٦ #(٢)	٨
المجموع	١٥٠	٥٠	٢٠٠

← (قع) أقل من ٥

في المثال (٣ , ٧) سميت خلايا جدول التوافق ٢ × ٢ بالأرقام الموضوعة داخل الدوائر بينما تركت التكرارات الملاحظة بالخلايا الأربع كما هي ولكن وضعت

تكراراتها المتوقعة بين قوسين. ولقد تم حساب التكرارات المتوقعة بتطبيق القاعدة (٣٤) كالمعتاد. ونلاحظ في هذا المثال أن هناك ثلاثة تكرارات متوقعة، وبالتحديد تكرارات الخلايا ②، ③ و ④، يقل كل واحد منها عن ١٠ مما يتطلب منا التدخل لتصحيح الوضع. ذلك أنه، وبترك الوضع على ما هو عليه، نتحصل على كا^{٢} محسوبة بمقدار ١٣,٥، بينما نقرأ من الجدول النظري لتوزيع كا^{٢} قيمة لـ كا^{٢} (الجدولية) بدرجة حرية «١» ومستوى دلالة ٠,٠٥ تساوي ٣,٨٤ مما يستدعي رفض الفرضية الصفرية المصاحبة للدراسة خاصة بيانات جدول المثال (٧,٣). ولتصحيح الوضع في المثال (٧,٣)، أي لتعديل قيمة كا^{٢} المحسوبة المتأثرة بالبسط في العملية الحسابية

$\frac{(ل - قع)^{٢}}{قع}$ ، اتفق على تطبيق ما يسمى بتصحيح «ياتس» Yates's correction

المعروف في مجال تصحيح مكونات قيمة كا^{٢} في الجدول التكراري ٢ × ٢ الذي تقل قيمة إحدى أو بعض التكرارات المتوقعة لخلاياه عن ١٠. ويصار هنا إلى طرح ٠,٥٠

من القيمة المطلقة لبسط العملية الحسابية $\frac{(ل - قع)^{٢}}{قع}$ قبل تربيعه ومن ثم إتمام إجراء

العملية. فمثلا إذا أردنا أن نجري تصحيح ياتس على الخلية ③ بجدول المثال (٧,٣) فسيكون ذلك كما يلي:

$$\text{كا}^{٢} \text{ (المحسوبة) الخاصة بالخلية ③ } = \frac{(|٦ - ٩,٣٣| - ٠,٥٠)^{٢}}{٩,٣٣}$$

$$= ٠,٨٥٨$$

مع ملاحظة أن الخطين العموديين اللذين يحصران الرقمين ٦ و٩,٣٣ يشيران إلى أننا نأخذ بالقيمة المطلقة (أي نهمل إشارة السالب) لناتج عملية طرح هذين الرقمين قبل أن نطرح ٠,٥٠ من هذا الناتج. وبغياب إجراء عملية التصحيح هذه، تكون قيمة كا^{٢} (المحسوبة) الخاصة بالخلية كما يلي:

$$\text{كا}^2 \text{ (المحسوبة) الخاصة بالخلية ③ } = \frac{(٦ - ٩,٣٣)^2}{٩,٣٣}$$

$$= ١,١٨٩$$

لاحظ أن الفرق بين القيمتين (المصححة وغير المصححة) يناهز ٢٨٪[٦] من القيمة غير المصححة، أي أن القيمة المصححة تقل بأكثر من ربع القيمة غير المصححة.

وبتطبيق تصحيح ياتس على الخلايا ②، ③ و ④ من الجدول بالمثال (٣ , ٧) نتحصل على كا^2 (المحسوبة) = ٣,٧١ بدلا من قيمتها قبل التصحيح والتي تساوي ١٣,٥ . والمفارقة أنه بينما نرفض الفرض الصفري قبل التصحيح [حيث تكون قيمة كا^2 (الجدولية) بدرجة حرية «١» ومستوى دلالة ٠,٠٥ و«٣,٨٤»] نرجع إلى قبوله بعد التصحيح [قيمة كا^2 (الجدولية) تظل كما هي ٣,٨٤ فيما تتحول قيمة كا^2 (المحسوبة) إلى ٣,٧١].

وفي جدول المثال (٤ , ٧) نلاحظ أن هناك خليتين، وبالتحديد الخلية ⑥ والخلية ⑧، تقل تكراراتهما المتوقعة عن ٥ . ومع أنه لا يطبق تصحيح ياتس في الجداول ذات الحجم الأكبر من ٢ × ٢ إلا أن معالجة خاصة لهذا الوضع والأوضاع المماثلة يمكن أن تتم باستخدام دمج الخلايا المعيبة مع بعضها إذا كان هذا يتماشى مع المنطق. وقبل أن نسترسل في كيفية علاج المشكلة، دعنا نتأمل حجمها أولا من واقع بعض التكرارات المتوقعة بجدول المثال (٤ , ٧). فلو أجرينا العملية الحسابية $\frac{(\text{ل} - \text{قع})^2}{\text{قع}}$ للخلية ⑧ في هذا الجدول لتحصلنا على قيمة لـ كا^2 (المحسوبة) الخاصة بهذه الخلية تساوي ٨ [(٦-٢)^2 ÷ ٢ = ٨] وهذه القيمة وحدها تزيد عن قيمة كا^2 (الجدولية) لبيانات الجدول المعني بعدد ٣ درجات حرية ومستوى دلالة ٠,٠٥ حيث تكون قيمتها ٧,٨١٥! هذا، ناهيك عن حساب قيم كا^2 المحسوبة لبقية الخلايا السبع وجمعها مع قيمة كا^2 المحسوبة

---

(٦) توجد هكذا:

$$\frac{(١,١٨٩ - ٠,٨٥٨)}{١,١٨٩} \times ١٠٠ = ٢٧,٨٣٩ .$$

الخاصة بالخلية ⑧ ، وهي تساوي ٨ كما أشرنا من قبل ، لنتحصل على قيمة خرافية لـ كا² المحسوبة بالمقارنة مع كا² الجدولية .

أما فيما يتعلق بالتصحيح المنوط به تحويلنا إلى قيمة لـ كا² المحسوبة تكون أكثر معقولية فهو كما أشرنا سابقا يتعلق بإمكانية دمج قيم الخلايا ذات الفئات المتشابهة لخلق جدول جديد أصغر حجما وأقرب لأن يمكننا من الحصول على قيمة أكثر معقولية من القيمة السابقة (أي قبل التعديل) . فمثلا ، إذا كان التصنيف لفئات متغير السرقة على سبيل المثال اقتضى فرز مجموعات صفوف الجدول من أعلى إلى أسفل إلى : سرقة عادية ، ثم نهب ، ثم احتيال ، ثم تزوير ، فإنه يمكننا دمج التكرارين الملاحظين للفئتين ٥ و٧ ثم دمج التكرارين الملاحظين للفئتين ٦ و٨ كلا على حدة لنتحصل على جدول جديد ٢ × ٣ بتكرار ملاحظ يساوي ٩ في الخلية ⑥ وتكرار ملاحظ يساوي ١١ في الخلية ⑤ . ثم نقوم بعد ذلك بحساب كا² لكل الجدول لنتحصل على قيمة جديدة أقرب للواقع مع ملاحظة أننا خفضنا عدد درجات الحرية الآن إلى (٢-١) (٣-١) = ٢ مما يؤدي أيضا إلى ارتفاع قيمة كا² الجدولية ونحن نبقي على مستوى الدلالة كما هو . ويمكن للدارس أن يجري هذه العمليات قبل وبعد دمج الخلايا ليتأكد بنفسه من حال الوضعين ، القبلي والبعدي ، فيما يتعلق بقيم كا² المحسوبة والجدولية في الحالتين وتأثير ذلك بالتالي على القرار بالرفض أو القبول لفرض العدم .

٥.٣.٧ خاتمـــة

بالرغم من أن لتوزيع مربع كا² استخدامات أخرى في مواضع تتعلق بتطبيق اختبارات لا معلمية ، إلا أن كثافة استخدامه في حالة العينات المستقلة من قبل الباحثين الاجتماعيين استدعى التركيز عليه هنا في مجاله التقليدي . ومتطلبات تطبيقه في مجال العينات المستقلة يقتضي :

١ - الإقبال نحو المقارنة بين عينتين أو أكثر من العينات المستقلة ولهذا يتطلب الأمر الحصول على جدول يكون حجمه على الأقل ٢ × ٢ .

٢- بيانات اسمية أو رتبية (وإن تكن معظم بيانات ا لبحوث الاجتماعية من هذا النوع).

٣- اختيار عشوائي للعينات المستقلة.

٤- لا يجب أن تكون التكرارات المتوقعة صغيرة الحجم بصورة لافتة. وعلى الأخص، يتعين على الباحث أن يوظف تصحيح ياتس عندما يكون هناك أي تكرار متوقع يقل عن ١٠ في حالة الجداول التي بحجم ٢ × ٢، كما لا يجب أن يكون هناك تكرار متوقع يقل عن ٥ بأي حال من الأحوال في هذا النوع من الجداول. وفي حالة الجداول التي يفوق حجمها ٢ × ٢ لا توجد هناك قاعدة ثابتة فيما يتعلق بمسألة حد أدنى للتكرارات الملاحظة رغم أن هذا لا ينفي أنه يجب علينا التأكد من أن خلايا قليلة هي التي تحمل تكرارات تقل عن ٥ في هذا النوع من أنواع الجداول. وفي مثل هذه الحالات ينبغي محاولة دمج مثل هذه الخلايا بعضها مع بعض إن كان لا يتعارض هذا مع منطق الأشياء.

٥- ألا نهتم بشرط التوزيع الطبيعي للمجتمع المسحوبة منه العينة.

٤,٧ أسئلة

١- أجر اختبارا للدلالة الإحصائية مستخدما كا^{٢} لبيانات الجدول ٣ × ٣ التالي علما بأن مستوى الدلالة المختار هو ٠,٠٥.

١٥	١٠	٨
٩	١٠	١٢
١٢	٨	٩

٢- تم اختيار عينتين إحداهما من الطلاب الذكور وأخرى من الإناث ووجه إلى أفراد كل عينة السؤال التالي:
هل تعرف استخدام الحاسب الآلي؟ وكانت الحصيلة أن ١٥ من ٢٩ طالبا أجابوا بالإيجاب فيما أجابت ٢٠ طالبة بالإيجاب من جملة ٣٠ طالبة. اختبر فرض

العدم الذي مفاده أن التكرار النسبي للطلاب الذين يعرفون استخدام الحاسب الآلي هو نفسه التكرار النسبي للطالبات اللاتي يعرفن استخدام الحاسب الآلي .

٣- مطبقا تصحيح ياتس ، أجر اختبارا للدلالة الإحصائية مستخدما كا^2 لكل مسألة من المسائل التالية :

(أ)

١٥	٧
١٠	٦

(ب)

٨	١٢
١٠	٥

(جـ)

١٨	١٢
٣	٨

٧,٥ اصطلاحات ينبغي تذكرها

- مربع كاي (كا^2) .
- كا^2 (المحسوبة/ الجدولية) .
- مستوى الدلالة .
- درجة الحرية .
- التكرار الملاحظ/ التكرار المتوقع .
- الإحصاء اللامعلمي/ الإحصاء المعلمي .
- تصحيح ياتس .
- فرض العدم/ الفرضية الصفرية .
- فرض البحث .
- الفروق ذات الدلالة الإحصائية .
- الفروق الصُّدْفية .

المراجـــع

أولا: العربية

أبو شعر، عبدالرازق. *مبادئ الإحصاء*. الرياض: معهد الإدارة العامة، ١٩٨٢م.

أبو صالح، محمد صبحي؛ عوض، عدنان محمد. *مقدمة في الإحصاء*. نيويورك: دار جون وايلي وأبنائه، ١٩٨٣م.

بلالوك، هيوبرت م. *الإحصاء الاجتماعي*، الطبعة الثانية المنقحة (١٩٧٩م)، ترجمة محمد نور، عثمان الحسن ورضوان، سليمان محمد. القاهرة: مطابع الأهرام، ١٩٩٣م.

توفيق، عبدالجبار. *التحليل الإحصائي في البحوث التربوية والنفسية والاجتماعية*. الكويت: مؤسسة الكويت للتقدم العلمي، ١٩٨٢م (الطبعة الأولى).

سرحان، أحمد عبادة. *مقدمة في طرق التحليل الإحصائي*. القاهرة: دار المعارف، ١٩٧٤م.

الشربيني، زكريا. *الإحصاء اللابرامتري في العلوم النفسية والتربوية والاجتماعية*. القاهرة: مكتبة الأنجلو المصرية، ١٩٩٠م.

عوض، أحمد صفي الدين. *الإحصاء العام*. الرياض: دار العلوم للطباعة والنشر، الرياض، ١٩٨٤م.

منشورات مكتب التربية العربي لدول الخليج. *الإحصاء التربوي*، الجزء الأول، ١٩٨٠م.

منشورات مكتب التربية العربي لدول الخليج. *الإحصاء التربوي*، الجزء الثاني، ١٩٨١م.

هويل، بول ج. *المبادئ الأولية في الإحصاء*، الطبعة الخامسة المنقحة، ترجمة عبدالوهاب، بدرية شوقي والشربيني، محمد كامل. نيويورك: دار جون وايلي وأبنائه، ١٩٨٤م.

ثانيا: الأجنبية

Bartholomew, David J. *The Statistical Approach to Social Measurement.* Academic Press, Inc., 1996.

Blalock, Hubert M. *Social Statistics.* 2nd revised edition, McGraw-Hill, Inc., 1979.

First Year Study Guides. *Quantitative Methods.* London: BPP Publishing, 1993.

Hansen, M. H., Hurwitz, W. N., and Madow, W. G. *Sample Survey Methods and Theory.* New York, London, and Sydney: John Willey & Sons, Inc., 1953.

Kish, Lesli. *Survey Sampling.* New York, London, and Sydney: John Willey & Sons, Inc., 1965.

Knoke, David and Bohrnstedt, George W. *Statistics for Social Data Analysis,* second edition, F.E. Peacock Publishers, Inc., N.Y., 1988.

Levin, Jack and Fox, James Alan. *Elementary Statistics in Social Research.* 4th edition, Harpert & Row, Publishers, Inc., 1988.

توزيع كاي تربيع Distribution of χ^2

الاحتمال / درجة الحرية	٠,٩٩	٠,٩٨	٠,٩٥	٠,٩٠	٠,٨٠	٠,٧٠	٠,٥٠	٠,٣٠	٠,٢٠	٠,١٠	٠,٠٥	٠,٠٢	٠,٠١	٠,٠٠١
١	٠,٠١٥٧	٠,٠٦٢٨	٠,٠٣٩٣	٠,٠١٥٨	٠,٠٦٤٢	٠,١٤٨	٠,٤٥٥	١,٠٧٤	١,٦٤٢	٢,٧٠٦	٣,٨٤١	٥,٤١٢	٦,٦٣٥	١٠,٨٢٧
٢	٠,٠٢٠١	٠,٠٤٠٤	٠,١٠٣	٠,٢١١	٠,٤٤٦	٠,٧١٣	١,٣٨٦	٢,٤٠٨	٣,٢١٩	٤,٦٠٥	٥,٩٩١	٧,٨٢٤	٩,٢١٠	١٣,٨١٥
٣	٠,١١٥	٠,١٨٥	٠,٣٥٢	٠,٥٨٤	١,٠٠٥	١,٤٢٤	٢,٣٦٦	٣,٦٦٥	٤,٦٤٢	٦,٢٥١	٧,٨١٥	٩,٨٣٧	١١,٣١٤	١٦,٢٦٨
٤	٠,٢٩٧	٠,٤٢٩	٠,٧١١	١,٠٦٤	١,٦٤٩	٢,١٩٥	٣,٣٥٧	٤,٨٧٨	٥,٩٨٩	٧,٧٧٩	٩,٤٨٨	١١,٦٦٨	١٣,٢٧٧	١٨,٤٦٥
٥	٠,٥٥٤	٠,٧٥٢	١,١٤٥	١,٦١٠	٢,٣٤٣	٣,٠٠٠	٤,٣٥١	٦,٠٦٤	٧,٢٨٩	٩,٢٣٦	١١,٠٧٠	١٣,٣٨٨	١٥,٠٨٦	٢٠,٥١٧
٦	٠,٨٧٢	١,١٣٤	١,٦٣٥	٢,٢٠٤	٣,٠٧٠	٣,٨٢٨	٤,٣٤٨	٧,٢٣١	٨,٥٥٨	١٠,٦٤٥	١٢,٥٩٢	١٥,٠٣٣	١٦,٨١٢	٢٢,٤٥٧
٧	١,٢٣٩	١,٥٦٤	٢,١٦٧	٢,٨٣٣	٣,٨٢٢	٤,٦٧١	٦,٣٤٦	٨,٣٨٣	٩,٨٠٣	١٢,٠١٧	١٤,٠٦٧	١٦,٦٢٢	١٨,٤٧٥	٢٤,٣٢٢
٨	١,٦٤٦	٢,٠٣٢	٢,٧٣٣	٣,٤٩٠	٤,٥٩٤	٥,٥٢٧	٧,٣٤٤	٩,٥٢٤	١١,٠٣٠	١٣,٣٦٢	١٥,٥٠٧	١٨,١٦٨	٢٠,٠٩٠	٢٦,١٢٥
٩	٢,٠٨٨	٢,٥٣٢	٣,٣٢٥	٤,١٦٨	٥,٣٨٠	٦,٣٩٣	٨,٣٤٣	١٠,٦٥٦	١٢,٢٤٢	١٤,٦٨٤	١٦,٩١٩	١٩,٦٧٩	٢١,٦٦٦	٢٧,٨٧٧
١٠	٢,٥٥٨	٣,٠٥٩	٣,٩٤٠	٤,٨٦٥	٦,١٧٩	٧,١٧٩	٩,٣٤٢	١١,٧٨١	١٣,٤٤٢	١٥,٩٨٧	١٨,٣٠٧	٢١,١٦١	٢٣,٢٠٩	٢٩,٥٨٨
١١	٣,٠٥٣	٣,٦٠٩	٤,٥٧٥	٥,٥٧٨	٦,٩٨٩	٨,١٤٨	١٠,٣٤١	١٢,٨٩٩	١٤,٦٣١	١٧,٢٧٥	١٩,٦٧٥	٢٢,٦١٨	٢٤,٧٢٥	٣١,٢٦٤
١٢	٣,٥٧١	٤,١٧٨	٥,٢٢٦	٦,٣٠٤	٧,٨٠٧	٩,٠٣٤	١١,٣٤٠	١٤,٠١١	١٥,٨١٢	١٨,٥٤٩	٢١,٠٢٦	٢٤,٠٥٤	٢٦,٢١٧	٣٢,٩٠٩
١٣	٤,١٠٧	٤,٧٦٥	٥,٨٩٢	٧,٠٤٢	٨,٦٣٤	٩,٩٢٦	١٢,٣٤٠	١٥,١١٩	١٦,٩٨٥	١٩,٨١٢	٢٢,٣٦٢	٢٥,٤٧٢	٢٧,٦٨٨	٣٤,٥٢٨
١٤	٤,٦٦٠	٥,٣٦٨	٦,٥٧١	٧,٧٩٠	٩,٤٦٧	١٠,٨٢١	١٣,٣٣٩	١٦,٢٢٢	١٨,١٥١	٢١,٠٦٤	٢٣,٦٨٥	٢٦,٨٧٣	٢٩,١٤١	٣٦,١٢٣
١٥	٥,٢٢٩	٥,٩٨٥	٧,٢٦١	٨,٥٤٧	١٠,٣٠٧	١١,٧٢١	١٤,٣٣٩	١٧,٣٢٢	١٩,٣١١	٢٢,٣٠٧	٢٤,٩٩٦	٢٨,٢٥٩	٣٠,٥٧٨	٣٧,٦٩٧

ثبت المصطلحات

أولا: عربي - إنجليزي

أ

إحصاء استنتاجي	Inferential statistics
إحصاء وصفي	Descriptive statistics
اختبار الدلالة الإحصائية	Test of statistical significance
اختبارات الفروض	Tests of hypotheses
ارتباط	Correlation
أعمدة بسيطة	Bar graphs
أعمدة مجزأة	Superimposed bar graph
أعمدة مزدوجة	Pair bar graph
انحراف متوسط	Mean deviation
انحراف معياري	Standard deviation

ب

بيانات خام	Raw data

ت

Variance	تباين
Experiment	تجربة
Content analysis	تحليل المضمون
Questionnaire design	تصميم الاستبانة
Experimental design	تصميم التجارب

ث

Constant	ثابت
Reliability	ثبات

ج

Data collection	جمع البيانات

ح

Sample coverage	حصر بالعينة
Complete coverage	حصر شامل

خ

Strip mapping	خريطة الشريط
Polygon	خط بياني

د

Pie chart	دائرة بيانية

Standardized score	درجة معيارية

ص

Validity	صدق

ع

Presentation of data	عرض البيانات
Graphic presentation	عرض بياني
Tabulation	عرض جدولي
Statistics	علم الإحصاء
Sample	عينة
Probabilistic sample	عينة احتمالية
Accidental sample	عينة بالمصادفة
Quota sample	عينة حصصية
Simple random sample	عينة عشوائية بسيطة
Stratified random sample	عينة عشوائية طبقية
Clustered random sample	عينة عشوائية عنقودية
Systematic random sample	عينة عشوائية منتظمة
Purposive sample	عينة عمدية
Non probabilistic sample	عينة غير احتمالية

ف

Research hypothesis	فرض بحث

فرض عدم	Null hypothesis
فرضيات	Hypotheses

ق

قياس	Measurement
قياس اسمي	Nominal measurement
قياس ترتيبي	Ordinal measurement
قياس فاصل (فتري)	Interval measurement
قياس نسبي	Ration measurement

م

متغيرات	variables
متغيرات تابعة	Dependent variables
متغيرات مستقلة	Independent variables
مدى	Range
مراسلة بالبريد	Mailing method
مربع كاي	Chi-Square
مستويات الدلالة	Significance levels
مسح اجتماعي	Social survey
مصدر أولي	Primary source
مصدر ثانوي	Secondary source
معاملات الارتباط	Correlation coefficients
مقابلة شخصية	Personal interview

Measures of dispersion	مقاييس التشتت
Measures of central tendency	مقاييس النزعة المركزية
Participant observation	ملاحظة بالمشاركة
Mode	منوال

ن

Semi-interquartile range	نصف المدى الربيعي
Probability theory	نظرية الاحتمالات
Estimation theory	نظرية التقدير

و

Observation unit	وحدة المشاهدة
Arithmetic mean	وسط حسابي
Median	وسيط

ثانيا: إنجليزي – عربي

A

Accidental sample	عينة بالمصادفة
Arithmetic mean	وسط حسابي

B

Bar graphs	أعمدة بسيطة

C

Chi-Square	مربع كاي
Clustered random sample	عينة عشوائية عنقودية
Complete coverage	حصر شامل
Constant	ثابت
Content analysis	تحليل المضمون
Correlation	ارتباط
Correlation coefficients	معاملات الارتباط

D

Data collection	جمع البيانات
Dependent variables	متغيرات تابعة
Descriptive statistics	إحصاء وصفي

E

Estimation theory	نظرية التقدير

تجربة	Experiment
تصميم التجارب	Experimental design

G

عرض بياني	Graphic presentation

H

فرضيات	Hypotheses

I

متغيرات مستقلة	Independent variables
إحصاء استنتاجي	Inferential statistics
قياس فاصل (فتري)	Interval measurement

M

مراسلة بالبريد	Mailing method
انحراف متوسط	Mean deviation
قياس	Measurement
مقاييس النزعة المركزية	Measures of central tendency
مقاييس التشتت	Measures of dispersion
وسيط	Median
منوال	Mode

N

قياس اسمي	Nominal measurement

Non probabilistic sample	عينة غير احتمالية
Null hypothesis	فرض عدم

O

Observation unit	وحدة المشاهدة
Ordinal measurement	قياس ترتيبي

P

Pair bar graph	أعمدة مزدوجة
Participant observation	ملاحظة بالمشاركة
Personal interview	مقابلة شخصية
Pie chart	دائرة بيانية
Polygon	خط بياني
Presentation of data	عرض البيانات
Primary source	مصدر أولي
Probabilistic sample	عينة احتمالية
Probability theory	نظرية الاحتمالات
Purposive sample	عينة عمدية

Q

Questionnaire design	تصميم الاستبانة
Quota sample	عينة حصصية

R

Range	مدى

Ration measurement	قياس نسبي
Raw data	بيانات خام
Reliability	ثبات
Research hypothesis	فرض بحث

S

Sample	عينة
Sample coverage	حصر بالعينة
Secondary source	مصدر ثانوي
Semi-interquartile range	نصف المدى الربيعي
Significance levels	مستويات الدلالة
Simple random sample	عينة عشوائية بسيطة
Social survey	مسح اجتماعي
Standard deviation	انحراف معياري
Standardized score	درجة معيارية
Statistics	علم الإحصاء
Stratified random sample	عينة عشوائية طبقية
Strip mapping	خريطة الشريط
Superimposed bar graph	أعمدة مجزأة
Systematic random sample	عينة عشوائية منتظمة

T

Tabulation	عرض جدولي

اختبارات الفروض	Tests of hypotheses
اختبار الدلالة الإحصائية	Test of statistical significance

U

صدق	Validity
متغيرات	variables
تباين	Variance

كشاف الموضوعات

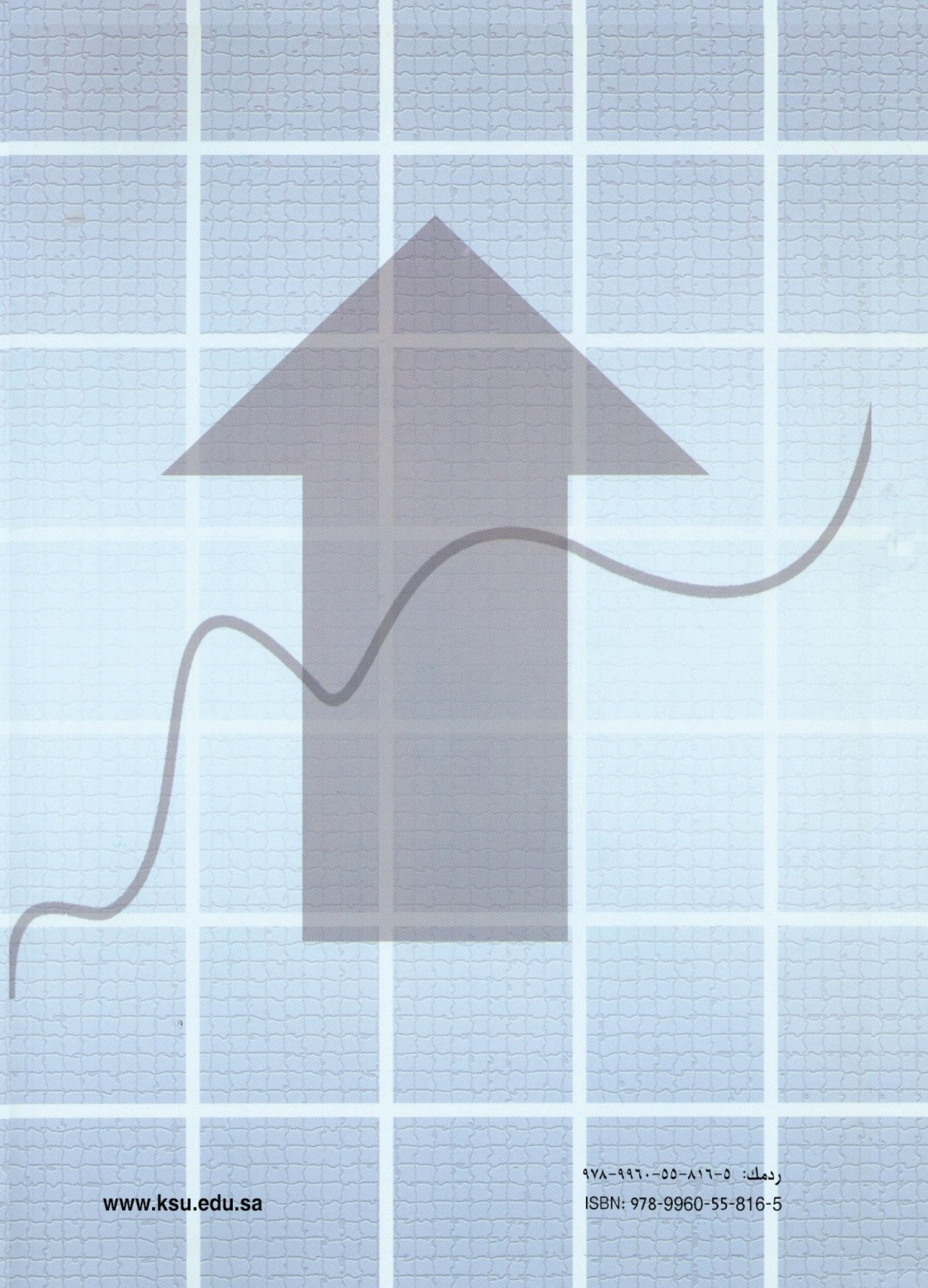
نبذة عن المؤلف

الدكتور/ صالح بن محمد الصغير

أستاذ وناقد في مجال الدراسات الاجتماعية (علم الاجتماع)، له اهتمامات في مجالات الإحصاءات الاجتماعية وعلوم المعرفة، والتغيرات الاجتماعية والتنمية والاتجاهات البيئية ودراسات استشراق المستقبل، وهو حاصل على بكالوريوس علم الاجتماع من جامعة الملك سعود بالرياض عام ١٩٨٦م، وماجستير علم الاجتماع من جامعة ولاية كلورادو عام ١٩٩٢م، ودكتوراه الفلسفة في علم الاجتماع والدراسات البيئية من جامعة ولاية المسيسيبي بالولايات المتحدة الأمريكية عام ١٩٩٥م، عمل أستاذا مساعدا في قسم الدراسات الاجتماعية في جامعة الملك سعود بالرياض منذ عام ١٩٩٥م، وأستاذا مشاركا منذ عام ١٩٩٨م، ورئيسا للقسم منذ عام ١٩٩٩م، كما عمل في عدد من المؤسسات والهيئات الحكومية، كما عمل عضوا في عدد من المجالس واللجان في جامعة الملك سعود ومازال. كما أشرف وناقش عدداً من رسائل الماجستير والدكتوراه داخل الجامعة وخارجها.

شارك في عدد من هيئات التحكيم لأبحاث الدراسات الاجتماعية، كما أنه عضو منظمة علم الاجتماع الأمريكية.

صدر له أكثر من (١٥) بحثا ومقالة علمية منشورة في مجال الدراسات الاجتماعية، يعمل حاليا رئيسا لقسم الدراسات الاجتماعية في كلية الآداب، جامعة الملك سعود، الرياض، المملكة العربية السعودية.



ردمك: ٥-٨١٦-٥٥-٩٩٦٠-٩٧٨
ISBN: 978-9960-55-816-5

www.ksu.edu.sa